1999

Evaluation of a Continuous Performance Test.

Sandi Louise Spera
Louisiana State University and Agricultural & Mechanical College

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

EVALUATION OF A CONTINUOUS PERFORMANCE TEST

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Psychology

by
Sandi L. Spera
B.A., University of New Orleans, 1990
M.A., Louisiana State University, 1994
May 1999

TABLE OF CONTENTS

LIST OF TABLES

ABSTRACT

Continuous performance tests are frequently used to measure attention and impulsivity in children and adults. These instruments are particularly popular for assessing Attention Deficit Hyperactivity Disorder. The ecological validity of continuous performance tests has not yet been established as acceptable (Barkley, 1991). Additionally, questions remain regarding the diagnostic utility of these instruments. This study examined the ecological validity and diagnostic utility of the Conners' Continuous Performance Test (CPT; Conners, 1995). Specifically, the relationship between CPT scores and behavior in a classroom setting was investigated. The ability of the Conners' Continuous Performance Test to discriminate between children who meet DSM-IV criteria for a diagnosis of Attention Deficit Hyperactivity Disorder and those who are classified as normal controls was also examined.

INTRODUCTION AND RATIONALE

Attention Deficit Hyperactivity Disorder (ADHD) accounts for approximately fifty percent of all school-age referrals to mental health clinics in the United States (Barkley, 1990). The prevalence of this disorder is estimated at 3% to 5% in school-aged children (APA, 1994). Essential features of the disorder are excessive inattention, impulsivity and hyperactivity. Thus, a comprehensive evaluation of ADHD includes assessment of functioning in each of these domains. One instrument often used to directly assess attention and impulsivity in a clinic or laboratory setting is the Continuous Performance Test (CPT; Rosvold, Mirsky, Sarason, Bransome & Beck 1956). Although originally designed to assess individuals with brain injury, various version of the CPT are now routinely included by many clinicians as part of a multi-method assessment of Attention Deficit Hyperactivity Disorder. Despite over forty years of clinical use, controversy regarding the ecological validity and diagnostic utility of these instruments exists. The literature is plagued with inconsistent findings and several important questions

regarding the usefulness of continuous performance tests have not yet been answered.

It presently remains unclear whether CPT scores actually represent levels of attention in a natural setting. Although, many studies have examined correlations between ratings of attention and CPT scores, relatively few studies have compared CPT scores with direct observations of on-task behavior in a classroom. This type of data is important as direct observations are a more ideal criteria for evaluating ecological validity than parent and/or teacher ratings (Barkley, 1991). In addition, classrooms are the setting where attention problems most often lead to concern. Importantly, none of the observation data reported in the literature was collected in a real classroom setting at a time when students were engaged in independent seatwork. Observations took place in lab playrooms, simulated classrooms, or on the playground. Additionally, the small number of studies that have included direct behavioral observations contain methodological problems which limit the degree to which results can be generalized. Small sample sizes and failure to address factors other than attention which may have

influenced CPT scores are two examples of the most common methodological flaws.

The degree to which the CPT can accurately discriminate children classified as Attention Deficit Hyperactivity Disordered from those classified as normal controls is also unclear. Several studies have calculated percentage agreement between classifications based on CPT scores and classifications based on other diagnostic instruments. While some researchers have reported percentage agreements of less than 55% (Gordon, DiNiro, Mettleman, 1988), others have reported percentage agreements as high as 70% (Fischer, Newby,& Gordon, 1995). Although a discriminate function analysis is appropriate for addressing concerns regarding the ability of continuous performance tests to accurately predict membership into diagnostic groups, few studies have utilized this statistical procedure to answer questions regarding the usefulness of continuous performance test for diagnosing ADHD. Studies that have utilized a discriminant function analysis yield conflicting findings and fail to provide sufficient information for readers to independently interpret reported findings. A discriminant function

analysis would be useful for examining the diagnostic utility of continuous performance tests alone and in combination with other measures of attention and impulsivity. This type of procedure would contribute valuable information to the current literature.

The proposed study has two primary purposes: (a) to evaluate the ecological validity of the CPT in the classroom and (b) to examine the diagnostic utility of continuous performance tests. Specifically, this study will attempt to answer the following empirical questions: 1) To what degree are CPT scores related to student behavior in the classroom as based on direct observation? 2) To what degree are CPT scores related to curriculum based measurement probes and ratings on the Behavior Evaluation Tool. 3) How well do CPT scores discriminate between groups of children who meet the DSM-IV criteria for ADHD and normal control children? 4) How well can CPT scores discriminate between the various sub categories of ADHD defined in DSM-IV? 5) How well can CPT scores discriminate between groups of children who meet the DSM-IV criteria for ADHD and normal control children as compared to other variables such as observations of behavior,

curriculum based measurement probes, and commonly used rating scales? 6) How well can CPT scores discriminate between groups of children who meet the DSM-IV criteria for ADHD and normal control children when combined with other variables such as observations of behavior, and/or curriculum based measurement probes? Preliminary psychometric data will also be collected on the Behavior Evaluation Tool.

LITERATURE REVIEW

Terminology

Before beginning a discussion of continuous performance tests, clarification of several important terms is necessary. The first of these is "attention". Attention is a construct that has been conceptualized and defined in numerous ways (Dember and Warm, 1979; Pick, Frankel and Hess; 1975). Although there is variability in the manner in which attention is operationalized, most researchers agree that attention is a broad term which refers collectively to numerous processes (Mesulum, 1985). Most broadly, attention is the concentration of mental effort on sensory or mental events (Solso, 1988). Attention is often discussed in terms of several different subtypes (Posner and Boies, 1971; Davies and Parasuraman, 1982; & Halperin, 1991). Typically, attention is divided into two major types; selective attention and sustained attention (See Halperin 1991 for a discussion). Selective attention refers to the ability to select from the environment those stimuli that are relevant for further processing. Sustained attention refers to the maintenance of information processing over time (Halperin, 1991).

Vigilance is a term that is related to attention, thus, it often appears in the literature on sustained attention. The term vigilance was first popularized in the 1950s by Norman Mackworth and was used to describe "a state of readiness to detect and respond to certain small changes occurring at random intervals in the environment" (Mackworth, 1950; Mackworth, 1957; Warm, 1984). Some researchers have postulated that, like attention, vigilance is a multidimensional concept, which refers to both a "level of vigilance" and a "vigilance decrement" (Corkum & Siegel, 1993). The term "level of vigilance" has been defined as an individual's overall ability to identify stimuli throughout the entire length of a task. "Vigilance decrement" is a term used to describe a decrease in an individual's ability to respond over time on task (Parasuraman & Davies, 1984).

History and Background of CPT

Human vigilance and sustained attention have been studied for several decades. Research in this area can be traced to the 1930's when there was concern regarding the ability of industrial workers to maintain accuracy in

detecting defective products after having worked on an assembly line for some time (Wyatt & Langdon, 1932). During World War II, there was a shift in the focus of this type of research from the performance of factory workers to the performance of military personnel. It was not until this time that the first series of controlled laboratory studies on human vigilance was conducted by Norman H. Mackworth to investigate the detection accuracy of radar operators who were required to attend to monotonous stimuli for long periods of time. Mackworth developed "the clock test" which allowed him to systematically observe and chart the vigilance of his subjects. He used the clock test to confirmed previous findings which suggested that the detection accuracy of radar operators decreased with an increase in time on watch. This phenomena became know as the "decrement function" or the "vigilance decrement" and it has been repeatedly demonstrated over the years (Dember & Warm, 1979; Davies & Parasuraman, 1982). Mackworth's research led to numerous studies in the areas of vigilance and attention.

Early studies in vigilance were the basis for the development of the first Continuous Performance Test by

Rosvold, Mirsky, Sarason, Bransome, and Beck in 1956. The first CPT consisted of a revolving drum on which two series of letters were mounted side by side. The drum revolved slowly revealing, in a window, one letter at a time. Subjects were asked to complete two ten minute tasks. The first task required the subject to press a key when he or she viewed an X in the window. The second task required the subject to press a key when he or she viewed an A followed by an X. The instrument was used to determine whether individuals with brain damage demonstrated deficits in sustained attention as compared to individuals who were not brain damaged (Rosvold et al, 1956). Rosvold and his colleagues proved the CPT useful in illustrating differences in the sustained attention of these two groups but noted that these differences may be due to factors other than impaired attention.

Throughout the seventies, studies emerged demonstrating that hyperactive children also performed poorly on vigilance tests (Campbell, Douglas, & Morgenstern, 1971; Sykes, Douglas, & Morgenstern, 1973). Since this time, the CPT has gained popularity as an instrument used to investigate attentional deficits in

children and adults with a variety of clinical disorders. Continuous performance tests have been strongly recommended by some researchers as an integral part of a multi method assessment battery for ADHD (Gordon & Mettleman, 1988).

Although there are several versions presently available, the modern versions of the CPT are not very different from the original version developed in 1956. Today, the continuous performance tests typically consists of stimuli (usually letters or numbers) which are repeatedly presented to a subject on a computer monitor. The subject is required to press a button when a certain stimulus or series of stimuli appears and to refrain from pressing the button when non-target stimuli appear. The CPT usually lasts between several minutes and one half hour. In addition to length of task, the various versions of CPTs vary on dimensions such as type of stimuli, modality of stimuli (visual or auditory), nature of task, interstimulus interval, number of targets, and display time.

Interpretation of CPT

Performance on the CPT is usually interpreted using three different scores; the omission score, the commission

score, and the total score. The omission score reflects the number of times a subject fails to respond to the target stimuli. Failure to respond is typically interpreted as a measure of inattention (Sostek, Buchsbaum, & Rapoport, 1980). Some researchers prefer to combine "very long latency responses" with misses to yield a more reliable and valid measure of inattention (Halperin, Sharma, Greenblatt, & Schwartz, 1991; Conners, 1995). The commission score indicates the number of times a subject responded to non-target stimuli. A commission score is frequently interpreted as a measure of impulsivity (Sostek et al, 1980). Further interpretation is possible by analyzing information on reaction time, reaction time variability, and anticipatory errors.

A less common form of interpretation involves using a CPT with a changing interstimulus interval. Specifically, the interstimulus interval varies as a function of performance. The interstimulus interval actually gets shorter when a correct response is made and gets longer following an incorrect response. Interpretation of results involves an examination of the changes in interstimulus intervals (Halperin, 1991).

Some investigators believe that it is necessary to interpret CPT performance using signal detection theory which involves examining measures of attentiveness or perceptual sensitivity (d') and the degree of caution in reporting an event as a signal or response bias (Beta) (Corkum & Siegel, 1993; Egan, Greenberg, & Schulman, 1961; Green & Swets, 1966; O' Dougherty, Nuechterlein, & Drew, 1984; Sostek, Buchsbaum, & Rapoport, 1980). Signal detection theory allows analysis of the level of vigilance, as well as, the separate analysis of the various components of vigilance (overall signal discrimination capacity, sustained attentional capacity over time, and inhibition of responses to irrelevant stimuli) (O'Dougherty, Nuechterlein, & Drew, 1984). The basis for using signal detection theory is the belief that vigilance decrements imply a shift to a more conservative response criterion rather than a decline in the observer's alertness or sensitivity during CPT tests (Warm, 1984). Signal detection theory, thus, allows one to examine such changes in response criterion as a function of time spent on task.

Reliability

Continuous performance tasks have been found to provide fairly reliable results across repeated trails. Gordon & Mettleman (1988) reported test retest correlations for the Gordon Diagnostic System (GDS; Gordon, 1983), one of the most frequently used continuous performance tasks. The GDS allows for administration of three basic types of tasks (two vigilance tasks and a "delay" task designed to measure impulsivity). A test-retest correlation of .72 was reported for total correct responses on the vigilance task when ninety children were retested between two and twenty two days following the first administration. This correlation dropped to .68 for total correct responses when children were tested one year later. Both correlations were reportedly significant at $p < .001$ level.

Greenberg examined test-retest reliability for another widely used continuous performance task, Test of Variables of Attention (T.O.V.A.)(Greenberg & Waldman, 1993). He retested thirty three subjects and reported that test-retest correlations were .5 or greater for commission errors on the TOVA. He also reported test-retest correlations of greater than .8 for mean and standard

deviation of response times. The test retest correlation for omission errors was only .14. Greenberg explained that this low number may have been due to the small number of omission errors made by children over age seven.

Variables Affecting Vigilance Performance

Age. The literature is rich with studies that have investigated variables influencing vigilance task performance. It appears that one's performance on a vigilance task may vary as a function of several different situational and task variables (Corkum & Siegel, 1993). One variable thought to affect vigilance performance is age. Many studies have reported a significant relationship between age and CPT scores (e.g. Fischer, Barkley, Edelbrock & Smallish, 1990; Halperin, Sharma, Greenblatt, & Schwartz, 1991; Hooks, Milich, & Lorch, 1994; Klee & Garfinkel, 1983; Lassiter et al., 1994; O'Dougherty, Nuechterlein, & Drew, 1984; Seidel & Joschko, 1990). It seems that the number of correct responses on vigilance tasks improves with increasing chronological age in children (See Warm, 1984 for a review). Interestingly, the greatest period of improvement appears to occur around the age of 8 or 9 years (Sykes, Douglas, Weiss & Minde, 1971).

Also, vigilance decrements have been shown to vary systematically as a function of age with more rapid declines occurring at higher chronological ages in groups of elementary and adolescent students (Kirby et al., 1978). A somewhat contrary finding is that older children were reported to be less susceptible than younger children to declines in their performance over time (Seidel & Joschko, 1990).

One of the first large scale studies to collect normative data on the vigilance performance of children was conducted by Levy (1979). Levy demonstrated an age related decrease in the number of commission and omission errors by children three to seven years old. These findings are similar to those reported in more recent study conducted by Greenberg and Waldman (1993). These researchers presented developmental normative data for 775 children age six to sixteen who completed the Test of Variables of Attention (T.O.V.A., Greenberg, 1987), which is a visual Continuous Performance Test. Results showed that the mean percentage of both total omission errors and total commission errors decreased curvilinearly with age. This decreasing curvilinear trend was also shown for the mean reaction time

and reaction time variability. Based on these results, the authors concluded that attention and impulse control develop in a non-linear manner, with rapid changes early in childhood and a "leveling off" during later childhood and early adolescence (Greenberg & Waldman, 1993). These results are different from those of Levy in that although she observed a decreasing trend in errors as chronological age increased, she did not observe the "leveling off" phenomena. It is possible that had Levy sampled children older than seven years of age, such an effect may have been demonstrated.

Although developmental trends in the vigilance performance of younger and middle age children seem apparent, there is still doubt as to the existence of age related differences in the vigilance performance of older teens and adults. Several attempts to study vigilance in this age group have yielded conflicting results (Neal & Pearson, 1966; Griew & Davies, 1962; Talland, 1966; Davies & Davies, 1975). However, after reviewing the literature in this area, Davies and Parasuraman (1982) concluded that older adults, particularly those over 60 years of age, perform worse than younger adults under certain task

conditions. For example, younger adult subjects appear to perform better than older adult subjects when the detection of more than one signal was required, when increased memory load is required, or when the event rate was high (Davies & Parasuraman, 1982). It is not clear if these findings are due to group differences in attention, perceptual sensitivity or group differences in decision criteria.

Gender. Another factor which has been repeatedly studied is gender. Davies and Tune (1969) reviewed vigilance studies conducted prior to 1969. Although these researchers noted numerous inconsistencies in the studies reviewed, they concluded that monitoring efficiency does not appear to be related to gender. More recent experiments yield conflicting results. While there is some evidence that sex-related differences in vigilance performance do not exist in either children or adults (Levy, 1979; Parasuraman, 1976; Sykes, Douglas, & Morgenstern, 1972; Kirchner & Knopf, 1974), there is also evidence that males make more omission, commission and anticipatory errors and have faster mean reaction times than females (Greenberg & Waldman, 1993). In one study, Horn, Wagner, and Ialong (1989) examined sex differences

using a subject pool of fifty four children diagnosed with Attention Deficit Hyperactivity Disorder and thirty one normal controls. Multivariate analyses of variance yielded no significance multivariate main effect for sex or for sex by diagnosis interaction.

Gordon and Mettleman (1988) conducted a larger study in which 1266 children, ranging in age from four to sixteen years, were administered a series of continuous performance tests. These included the Delay Task, the Vigilance task, and the Distractibility task, each of which can be administered with the Gordon Diagnostic System (Gordon, 1983). Several dependent measures were collected for each task. Significant main effects for sex surfaced for at least one dependent measure on each of the three tasks. The significant effects, however, were not consistent across age groups. Most importantly, the authors noted that sex accounted for no more than 2% of the variance within any of the age groups.

Overall, it appears more research is necessary before any valid conclusions regarding the effects of gender on vigilance performance can be made. Regardless, some

publishers of continuous performance tests suggest using norms based on both age and sex to interpret results.

Intelligence. Several studies have investigated the possibility that vigilance performance is related to level of intelligence. When comparing the monitoring behavior of children whose IQ scores fall within the average range, the majority of the literature suggests that no significant differences exist (McGrath, 1960; Wilkenson, 1961; Halcomb & Kirk, 1965; Kupitz & Richardson, 1978; Margolis, 1973; Gordon, 1988). The same is true for the vigilance performance of adults (Halcomb & Kirk, 1965; Ware, 1961). However, results from a recent study conducted by Swanson and Cooney (1989) imply that there is a moderate correlation between children's verbal IQ and their performance on a CPT. These authors used signal detection to derive four separate measures of vigilance performance for sixty three children in grades five through seven. Not only did they report significant correlations between the vigilance measures and IQ, but they reported significant correlations between vigilance performance and level of achievement as well.

When comparing the vigilance performance of individuals whose cognitive functioning is measured to be within the mentally disabled range, to individuals of average intelligence, a few studies have failed to demonstrate significant differences between groups (Kirby, Nettelbeck & Bullock, 1978; Ware, Baker, and Sipowicz, 1962). However, the results of several other studies have suggested that those classified as mentally disabled perform less well than those whose intellectual functioning is within the average range (Warm, 1984). Interestingly, Tomporowski and Allison (1988) found no significant differences between developmentally disabled and non-developmentally disabled individuals on one particular vigilance task, but found that non-developmentally disabled individuals performed better than developmentally disabled individuals on a different vigilance task which required the use of memory. Another group of researchers have postulated that observed differences in the vigilance performance of mentally retarded and non-retarded individuals is not the result of a true attentional deficit but is instead due to a developmental lag (Stanovich, 1978). Kirby, Nettelbeck & Thomas (1979) found that mildly

mentally retarded children showed an earlier and more rapid decline in vigilance performance than a control group of the same chronological age, however, the retarded children performed similarly to a control group with an equivalent mental age. These findings provide support for Stanovich's theory of developmental lag.

SES. Differences in socioeconomic status (SES) also appear to result in differences in vigilance performance. Three studies have demonstrated that children belonging to a low SES group performed worse on vigilance tests than did children from a high SES group (Mabel, 1968; Knopf and Mabel, 1975; Levy, 1980). Interestingly, the study conducted by Levy demonstrated that children from low SES backgrounds were approximately six months delayed in their ability to complete one particular version of the CPT as compared to same age peers from families classified as being of high SES (Levy, 1990). In other words, the performance of children from the low SES group was equivalent to that of children in the high SES group who were six months younger. No hypothesis were generated regarding the cause of these differences. To date, minimal research exists on SES and vigilance performance. More

research is necessary before definitive conclusions regarding the effect of SES on CPT performance can be made.

Task Variables. As noted previously, there are numerous versions of continuous performance tasks and most of these vary with regard to task variables. Often, little attention is paid to the variations in task variables and/or the effects of various task variables on performance. Display time and inter-stimulus interval (ISI) are two variables which have been found to affect performance on vigilance tasks. Studies have repeatedly shown that longer display times result in better performance on the CPT than shorter display times (Chee, Logan Schachar, Lindsay, & Wachsmuth, 1989). Increases in display time appear to reduce vigilance decrements while decreases in display time seem to increase the vigilance decrements and reduce the overall efficiency (Davies & Parasuraman, 1981). Corkum and Siegel (1993) reviewed several studies investigating the CPT performance of ADHD children and non ADHD controls. They reported that a stimulus exposure duration of 50 to 200 milliseconds seemed to best differentiate ADHD from non-ADHD control subjects. Studies that employed stimulus exposure durations longer

than 50 to 200 milliseconds tended to find smaller differences between ADHD and control groups.

The literature regarding ISI, unlike that regarding display time, is not very clear. It has been suggested that longer ISIs result in better performance. Sykes et al. (1971) manipulated ISI and reported that hyperactive, as well as, control children made more correct responses and fewer false alarms when the ISI was 1.5 seconds as compared to an ISI of 1.0 seconds. Interestingly, 1.5 seconds was found to be the most commonly used ISI in the literature involving ADHD children (Corkum & Siegel, 1993). Chee, Logan, Schachar, Lindsay and Wachsmuth (1989) manipulated display time and stimulus onset asynchrony (SOA). SOA is similar to interstimulus interval in that it is the length of time from the onset of one stimulus to the onset of the next stimulus. Chee et al.(1989)reported that subjects, regardless of group assignment (ADDH, CD, ADDH/CD, or LD, Control), missed less targets when the event rate was 2 seconds as compared to either a 1 second or a 4 second SOA. Interestingly, the researchers stated that the hit rate for the ADHD group deteriorated significantly faster at the fast and slow event rates than did the hit rate for the

other diagnostic groups. Chee et al.(1989) also reported that false alarm rate and reaction time was significantly higher for all groups when the event rate was slow as compared to a fast event rate. The researchers explained the findings with the hypothesis that subjects perform worse when the event rate is so low that they cannot finish attending and responding to one stimulus before the next appears. The researchers explained that as interstimulus interval increases there is time for attention to wander between trials, thus causing performance to deteriorate at longer SOAs when attention is off task.

Time on task is another variable which may be important to consider when evaluating CPT performance. Parasuraman and Davies (1984) suggest that vigilance decrements only occur after 30-45 minutes of time on task. However, decrements have been reported after as little as 2-5 minutes (Neuchterlein 1983). Corkum and Siegel (1993) reviewed thirteen studies examining the CPT performance of ADHD children and reported that time on task did not appear to be related to performance. The authors found that studies which required subjects to attend to task for a longer time period were no more likely to find differences

between the groups than studies which required subjects to attend to task for a shorter time. The authors pointed out, however, that the results must be interpreted with caution as other variables may have been responsible for the reported findings in the studies reviewed.

Chee et al.(1989) hypothesized that the variable time on task may interact with other variables to affect subject performance. Chee and her colleagues conducted a follow up study to the earlier one which examined display time and ISI. The team was interested in time on task as a confounding variable. It was reported that time on task did not appear to confound the effects of event rate on reaction time and false alarm rates, however, time on task may have confounded the effects of event rate on the hit rates observed in the previous study.

One final variable that may affect performance on continuous performance test is the nature of the task (Schachar et al., 1988). Presently, there are two versions of the CPT which are most commonly employed. One version, also termed a successive discrimination task, requires the subject to respond to a particular target only if it is preceded by another particular stimuli. For example,

subjects may be asked to respond to the letter X only if it is preceded by the letter A. In the other version, subjects are asked to respond to a specific target every time it is presented. For example, subjects may be asked to respond each time the letter X appears. After reviewing the literature, Parasuraman and Davies (1984) concluded that vigilance decrements occurred only in studies using successive discrimination tasks. However, in a recent review of 13 studies that examined the CPT performance of ADHD children, Corkum and Siegel (1993) concluded that studies were as likely to find significant differences between groups regardless of the type of task. One hypothesis for the conflicting findings may be that Corkum and Siegel's review centered on studies involving hyperactive children. It is possible that the performance of hyperactive children deviates so significantly from the norm that nature of task effects are masked. To date, this hypothesis has not been empirically investigated. Thus, the effects of the type of task remains equivocal.

After reviewing the literature regarding situational and task variables which affect performance on continuous performance tests, several conclusions can be made. First,

it appears that CPT performance improves with age in children. There is still doubt, however, as to whether older adults perform better than younger adults on vigilance tasks. Second, the effects of gender, socioeconomic status, and intelligence are unclear at this time. The majority of the literature suggest that these factors do not have a significant impact on vigilance task performance. Third, display time and interstimulus interval, have been demonstrated to influence CPT performance. Individuals have been found to perform better on continuous performance tasks when the stimulus duration or display time is long. Finally, faster events rates seem to be associated with shorter reaction times and lower false alarm rates (shorter ISIs). Additionally, there seems to be an optimal interstimulus interval which is neither the slowest nor fastest interval.

External Variables. External variables have also been found to been found to influence vigilance performance (See Corkum & Siegel, 1993 for a review). For example, Drager, Prior and Sanson (1986) demonstrated that vigilance decrements on an auditory and visual attention task were greater in a group of hyperactive children as compared to a

group of non-hyperactive control subjects when the experimenter was absent from the room. There was no significant difference between the groups when the experimenter remained in the room. This finding led the researchers to conclude that hyperactive children possessed a deficit in "application" rather than "ability". Critiques have suggested that the error rates on the CPT in this study were extremely low and the instrument was therefore insensitive to the attentional deficits of the ADHD children (Barkley, 1991). Although this study raises some interesting questions, it is the only one of its kind to date.

Surprisingly, the way in which an experimenter introduces a task is also likely to effect performance on a vigilance task. For example, when speed is emphasized, fast but less accurate responses are likely, however, when accuracy is emphasized, slower but more accurate responses are likely (Sergeant & Scholten, 1985). Interestingly, this effect was demonstrated in two non-hyperactive groups of children but not in a group of hyperactive youngsters (Sergeant & Scholten, 1985).

Another external variable is performance feedback or knowledge of results. Early experiments have demonstrated that performance feedback results in an increase in the frequency and speed of signal detection as well as a decrease in false responses (Chinn & Alluisi, 1964; Mackworth, 1950; McCormack, 1959; Loeb and Schmidt, 1960; Warm et al. 1973). A recent study conducted by O'Dougherty, Nuechterlein, and Drew (1984) replicated this finding. This study revealed that feedback (the ringing of a bell following correct responses) increased the overall hit rate as well as the overall perceptual sensitivity level of three groups of children classified as either hyperactive, hypoxic, or normal controls. The informational feedback did not produce differential effects on the overall performance of the three groups.

Chinn & Alluisi (1964), demonstrated that the effect of providing a specific type of information tended to be specific to the measure of performance efficiency used. For example, when information regarding missed-signals was provided, there was a significant decrease in the total number of false responses and when information regarding correct-detection was provided there subjects exhibited a

significant decrease in the over-all proportion of missed signals. Interestingly, knowledge of results for both omissions and commissions have been demonstrated to improve correct detection more than knowledge of results for either one of the variables alone (Mackworth, 1970). Also, performance feedback provided during a training task has been demonstrated to result in improved performance during vigilance tasks administered at a later time (Wiener, 1963).

Additionally, in 1972, Warm et. al. demonstrated that subjects who received feedback from the experimenter regarding response latency and subjects who engaged in self evaluation of response latency demonstrated enhanced performance on a vigilance task as compared to a control group that received no feedback and did not self evaluate. The researchers noted that the performance of the self-evaluation group was not significantly different from that of the experimenter-feedback group even though the self-evaluations were less accurate than the experimenter feedback.

Other studies have also revealed interesting findings. Not only have improvements in vigilance tasks been

demonstrated when subjects have been provided with feedback of limited accuracy, but improvements in vigilance performance have also been exhibited when subjects were given false or inaccurate feedback (Mackworth, 1964; Antonelli and Karas, 1967; Weidenfeller, Baker, and Ware, 1962).

In 1962, Sipowicz, Ware and Baker attempted to gather information regarding the effects of both knowledge of results and rewards on vigilance task performance. They investigated the effects of performance feedback and rewards both singly and in combination. Subjects were assigned to either a control group, reward group, performance feedback group (KR), or performance feedback and reward group. Subjects in the reward group received money if they detected all signals correctly but lost money if they made errors. They were given no performance feedback during the testing session. Subjects in the KR group were informed of all missed signals through the use of a lamp which was illuminated when a signal was missed. Subjects in the reward and performance feedback group received money following the session for good performance and at the close of the session lost money contingent on

missed signals. Subjects were also informed of missed signals during the session via a lamp. All experimental groups performed significantly better than the control groups, however, the group receiving a combination of reward and performance feedback exhibited the highest level of signal detection.

Reinforcement and Continuous Performance Tests

The results of early performance feedback studies lead researchers to hypothesize that providing subjects with knowledge of results improves performance because it reinforces appropriate responding on vigilance tasks (Dember & Warm, 1979; Warm, 1984). For example, Chinn & Alluisi (1964) used the concepts of positive and negative reinforcement to explain their finding that verbal feedback caused changes in CPT responses. They suggested that negative reinforcement was provided in the form of feedback regarding false responses. Correct responses were described as being positively reinforced by positive feedback.

A second hypothesis that has also received attention purports that knowledge of results plays an informational role in vigilance tasks by fostering observer's awareness of task characteristics (Adams & Humes, 1963; Baker, 1963).

In 1971, Warm et al., attempted to more closely investigate whether performance feedback actually reinforced vigilance performance. The researchers postulated that if knowledge of results actually served as a reinforcer for responding during vigilance tasks, then other principals of reinforcement would also apply to responding during vigilance tasks. Specifically, the investigators attempted to determine whether intermittent reinforcement leads to more stable levels of responding when reinforcement is subsequently withdrawn than does exposure to continuous reinforcement. Such an effect has been termed the partial reinforcement effect (Cofer and Appley, 1964). During the Warm et al.(1971) experiment, observers watched for illuminations of a lamp for one hour. Five different feedback conditions were arranged. Findings revealed that both the partial feedback schedule group and continuous feedback schedule group had similar response times and no vigilance decrements. However, when feedback was withdrawn during the last portion of the watch, the magnitude of vigilance decrements was greater for observers provided with continuous feedback than for observers provided with

intermittent feedback during the session. Thus, the hypothesis of the researchers was supported.

Only a handful of other researchers have pursued investigations in the area of reinforcement and vigilance. In Chinn & Alluisi's 1964 study, the researchers explained their findings using negative and positive reinforcement. For example, it was explained that negative reinforcement was presented in the form of feedback regarding false responses. Correct responses were described as being positively reinforced with positive feedback.

Most studies examining the effects of reinforcement on vigilance performance were conducted over two decades ago using early models of vigilance tests. No studies involving today's versions of continuous performance tests have been published. If performance on a continuous performance test can be manipulated with reinforcement, interesting implications regarding attentional deficits exist.

Continuous Performance Tests and ADHD

As previously discussed, continuous performance tests are more frequently being administered in research and clinic settings as part of a multi-method assessment battery for diagnosing Attention Deficit Disorder. The CPT

is one of few measures which assess both attention and impulsivity; two of the three hallmark characteristics of the disorder. Barkley (1991) purports that the CPT provides valuable information when assessing attention and impulsivity and suggests using the CPT as part of a multi-method assessment protocol for Attention Deficit Hyperactivity Disorder. While the CPT appears to provide a convenient and quick assessment, one might question the degree to which the information gathered using the CPT in a clinic represents levels of attention and impulsivity in the natural setting (e.g. home or school). Also, one may question whether information gathered using the CPT can be used to plan interventions for improving attention and impulsivity in the classroom or home setting. In other words, one might speculate about whether the CPT possess adequate ecological validity.

Ecological validity is a term that most often refers to the degree to which the results of a laboratory measure represent the actual behaviors of interest as they occur in naturalistic settings (Barkley, 1991). This type of validity has been more traditionally referred to as concurrent or predictive validity (Anastasi, 1976). It is

different from construct validity which refers to the degree to which a test measures a particular construct of interest (Sattler, 1990). In an attempt to address questions regarding ecological validity, as well as other more specific concerns regarding the utility of CPTs with the ADHD population, it is useful to look carefully at the existing validity studies.

Validity

Behavior Rating Scales. Many of the studies assessing the validity of continuous performance tests for evaluating ADHD have attempted to determine if scores on these tests are related to other currently available measures of attention and impulsivity. Behavior rating scales are one of the most commonly used diagnostic measures of attention and impulsivity. Many studies have investigated the degree to which CPT scores correlate with behavior rating scales. Studies have examined CPT scores and ratings of both impulsivity and hyperactivity. Examinations of both parent ratings and teacher ratings have been conducted as well.

Of the studies investigating the relationship between parent ratings and CPT scores, most have investigated the relationship between parent ratings of hyperactivity and

CPT scores. This is somewhat curious because the CPT is not purported to measure hyperactive behavior. Only a few studies have investigated the relationship between performance on the CPT and parent ratings of attention and impulsivity.

In 1988, Pascaulvaca, Wolf, Healey, Tweedy, & Halperin reported moderate to high correlations between CPT omission errors and parent ratings of hyperactivity and impulsivity (.45 to .51) in a group of ADHD girls. It is important to note, however, that significant correlations were not found by the researchers in a comparison group of ADHD boys. In 1991, Barkley reported that in a group of 6 to 11 year old ADHD and normal control children, omission errors on a CPT (Gordon, 1983) were significantly correlated ($p<.0037$) with parent ratings on the Child Behavior Checklist Hyperactivity factor (.29) and the Impulsivity/ Hyperactive factor of the Conners Parent Rating Scale-Revised (.34). Commission errors for the same group of children were also significantly correlated ($p<.0037$) with parent ratings on the Child Behavior Checklist Hyperactivity factor (.22) and the Impulsivity/Hyperactive factor (.34) of the Conners Parent Rating Scale.

During another study, Barkley gathered information from a group of ADHD and normal adolescent subjects ages 12 to 20 (Barkley, 1991). The omission scores from this group of children did not correlate significantly with parent ratings on the Child Behavior Checklist Hyperactivity factor or the Impulsivity/Hyperactive factor of the Conners Parent Rating Scale-Revised. However, the commission scores did correlate significantly (.36 and .25) with parent ratings on the above mentioned factors of the CBCL and the CPRS-R. Lassiter, D'Amato, Raggio, Whitten, & Bardos (1994) also reported significant correlations between CPT commission errors and scores on the Hyperactivity index of the Conners Parent Rating Scale.

In contrast, DuPaul et al. (1992) reported no significant correlations after examining CPT commission errors, CPT total correct scores, scores on the Home Situations Questionnaire (Barkley & Edelbrock, 1983), scores on the Hyperactivity scale of the Child Behavior Checklist (CBCL; Achenbach & Edelbrock, 1983), and scores on the Externalizing scale of the CBCL. The participants in this study were children diagnosed with ADHD. Lovejoy and Rasmussen (1990) also reported non-significant

correlations between parent ratings on the Hyperkinesis Index of the Conners Parent Rating Scale and CPT errors. The participants in this study consisted of one hundred elementary school children referred for learning and attentional difficulties.

Overall, limited and conflicting information exists on the relationship between parent ratings of behavior and scores on the CPT. While some evidence suggests that CPT scores are significantly correlated with ratings of attention, hyperactivity and impulsivity, other studies have failed to obtain significant correlations between these measures.

Many studies have also examined the relationship between teacher ratings and CPT scores. In reviewing these studies, it was difficult to summarize the literature due to the variety of rating scales and continuous performance tests examined, as well as, the conflicting results among various studies. Total error scores on CPTs have been reported to be significantly correlated with the inattention and hyperactivity factors of the Conners Teacher Rating Scale (Klee & Garfinkel, 1983) and teacher ratings on the ADHD Rating Scale (Wherry et al., 1993).

CPT omission scores have been noted to correlate significantly with the Inattention and Hyperactivity factors of the Conners Teacher Rating Scale (Klee & Garfinkel, 1983), teacher ratings on the ADHD Rating Scale (Wherry, Paal, Jolly, Adam, Holloway, Everett, and Vaught,1993), and the Inattention-Passivity factor of the Conners Teacher Questionnaire-28 (Halperin et al, 1988). Commission scores on the CPT have been reported to be significantly correlated with the Aggressivity and Hyperactivity factors of the Conners Behavior Rating Scale (Kupietz & Richardson, 1978), the Inattention and Hyperactivity factors of the Conners Teacher Rating Scale (Klee & Garfinkel, 1983) and scores on the Oppositional factor of the ADD-H Comprehensive Teacher's Rating Scale (Lassiter et al., 1994). Interestingly, Halperin et al (1988) examined the relationship of each of four types of commission errors and teacher ratings of behavior using the Conners Teacher's Questionnaire. Halperin et al.(1991) noted that only two of the four types of errors (e.g. A-not-X and X-only) were significantly correlated with teacher ratings.

Barkley (1991) looked at groups of older and younger children separately. He examined the relationship between teacher ratings of behavior and omission and commission scores in both elementary and adolescent groups of ADHD and non-ADHD children. Results differed for the two groups. Omission and Commission scores correlated significantly with teacher ratings on Conners hyperactivity and inattention index as well as the CAP overactivity and inattention index for 6-11 year olds. No significant correlations were noted within the 12-20 year old group.

Lovejoy & Ramussen (1990) also failed to report significant correlations between either omission or commission errors on the CPT (Lindgren & Lyons, 1984) and ratings on the IOWA Conners aggression and inattention /overactivity factors. Additionally, non-significant correlations were obtained by these researchers between CPT omission and commission scores and the Inattention-Passivity and Conduct Problem factors of the Conners Teacher Rating Scale. DuPaul et al (1992) also reported non significant correlations between the CPT scores of children diagnosed with ADHD and teacher ratings on several rating scales. Rating scales examined included the School

Situations Questionnaire (Barkley & Edelbrock, 1987), the Externalizing factor, Inattentive factor and Nervous-Overactive factor of the Child Behavior Checklist- Teacher Report Form (Achenbach & Edelbrock, 1986), and the ADHD Rating Scale (DuPaul, 1991). Significant correlations were not noted between any CPT scores and the Attention scale of the ADD-H Comprehensive Teacher's Rating Scale or the Hyperactivity scale of the same measure (ACTeRs; Ulmann, Sleator & Sprague, 1989; Lassiter, D'Amato, Raggio, Witten, and Bardos, 1994). Finally, Wherry et al. (1993), reported no significant correlations between any of the scores on the delay and distractibility tasks of the Gordon Diagnostic System and teacher ratings on the ADHD Rating Scale. McClure and Gordon (1984) later reported significant negative correlations between the number of responses made on the GDS delay task and the Achenbach Teacher Report form Hyperactivity index.

Overall, while there is some evidence of a statistically significant relationship between CPT scores and rating scale data obtained by teachers, there is also evidence that these measures are not significantly correlated.

There are several potential reasons for the conflicting results among studies investigating the relationship between parent and teacher ratings and CPT scores. First, factor analysis of rating scales may yield dimensions that are labeled as assessing a construct but are actually contaminated by items that do not truly represent the construct. The factors may instead represent the items that covary with it. It is possible that these covarying items reduce the maximum correlation that could be obtained between the true factor and scores on a CPT (Barkley, 1991). Second, the CPT may provide information about inattention and impulsivity that contributes to, but is not identical to the exhibition of these difficulties in natural settings (Halperin, 1991). Third, inattentive and impulsive behaviors may differ in a clinic setting due to the greater novelty of the situation (Barkley, 1990). Fourth, rating scale data is usually comprised of collapsed observations over an extended period of time, CPTs occur within only a few minutes (DuPaul, 1992). Finally, as previously discussed, various extraneous factors (e.g. age, situational factors, and external factors) may influence scores on CPT tests. Most of the studies reviewed failed

to address such factors, thus little information exists regarding whether or not these factors were controlled. It is possible that the conflicting results in the literature are due, at least in part, to the influence of one or several such factors.

Other Laboratory Measures of Attention and Impulsivity. In addition to being compared to behavior rating scales, CPT scores have also been examined in relation to scores obtained on other laboratory measures of attention and impulsivity. Again, the literature is difficult to summarize due to the various age groups, independent measures, and dependent measures utilized. Overall, it appears that scores on the CPT correlate significantly with the Freedom from Distractibility Factor of the WISC-R, which is purported by some to assess attention (Grant, Ilai, Nussbaum, & Bigler, 1990; Lovejoy & Rasmussen, 1990). Several investigators have also reported significant correlations between CPT scores and scores on the Matching Familiar Figures Test (McClure & Gordon, 1984; Barkley, 1991; Klee & Garfinkel, 1983). For example, Klee and Garfinkel (1983) reported that omission scores on a CPT correlated positively with total error scores on the MFFT

and negatively with latency of first response on the MFFT. Commission scores on a CPT were also noted to correlate positively with total error scores on the MFFT, as well as, correlate negatively with latency and total response time on the MFFT. Total error scores on the CPT were noted to correlate positively with total error scores on the MFFT and negatively with MFFT first response time. Overall, it should be noted, that the statistically significant correlations obtained between CPTs and the MFFT were of a low magnitude.

DuPaul et al. (1992) failed to find any significant correlations between scores on a CPT and scores on the MFFT in a group of children classified as ADHD (DuPaul et al, 1992). He also reported non-significant correlations between CPT scores and time to first response on the MFFT. Omission and commission scores were also not reported to correlate significantly with any of MFFT-20 scores in a group of ADHD and non-ADHD adolescents (Barkley, 1991).

Several attempts have been made to demonstrate significant correlations between measures of activity level (ankle and wrist actometers) and CPT performance. Because activity level measures are reportedly able to discriminate

between ADHD and normal children, a significant correlation between these measures and CPT scores would support the ecological validity of CPTs (Barkley, 1991; Barkley & Cunningham, 1979; Luk, 1985; Tryon, 1984). Barkley (1991) did not find wrist actometer scores to correlate significantly with the scores from a CPT. Ankle actometer score collected during the same study, however, did correlate significantly with commission scores.

Finally, correlations have been computed for CPT scores and scores obtained on the Children's Checking Task (CCT; Margolis, 1972). The CCT is a pencil and paper version of a continuous performance test which requires children to listen to numbers on a tape recorder while checking them against an almost identical series of numbers in a booklet. No significant relationships were reported between computerized CPTs and the pencil and paper versions (Lovejoy & Rasmussen, 1990).

In summary, there is limited support for the concurrent validity of CPTs as measured by their correlation with other measures of attention and impulsivity . Reported correlations between CPT scores and similar measures appear to be stronger within the

population of elementary aged children than within the population of adolescents and young adults. Overall, consistent and robust correlations between CPT scores and scores on other measures of attention and impulsivity have not been reported.

Halperin (1991) has pointed out several possible reasons why various measures of attention do not correlate well with each other. First, it is possible that these measures actually assess different aspects of attention. Second, each measure may assess a wide range of other cognitive functions which distinguishes it from the other measures. If this is true, then the common variance accounted for by attention may only amount to a small proportion of the total variance of these measures. Barkley, perhaps, best hypothesized the reason for the low correlations. In a summary of his own findings he stated "while the CPT scores share some variance with other laboratory measures of attention and impulsivity, they are hardly measuring identical constructs" (Barkley, 1991).

Direct Observations. Direct behavioral observations in natural settings have been noted by Barkley (1991) to be a better criteria for assessing ecological validity than

other measures such as behavior rating scales. Direct observations of behavior in natural settings are "high in validity and strong in their representativeness of natural behavior", thus falling at the far end of the ecological validity continuum (Barkley, 1991). Barkley (1991) cites several reasons why he considers direct observations to be more ideal criteria than rating scales. First, response categories or behaviors are better defined than items on rating scales. Second, direct observations are not as subject to error variance by factors that affect the informant rather than the subject of the ratings (e.g. maternal depression, marital discord etc.). Third, direct observations do not average out important situational or temporal fluctuations which may be essential to understanding the nature of the symptoms of ADHD and which may prove theoretically important in defining the disorder itself. Fourth, direct behavioral observations do not contain factor structures as do rating scales, thus, contaminants in factor structures and covariants are not a concern.

Despite their necessity, correlation studies aimed at examining CPT scores and observations of behavior in

natural settings are scarce and not promising. Although the CPT is routinely used to assess attentional difficulties in children who display inattention at school, surprisingly few studies have attempted to investigate the degree to which CPT results correlate with observations of on-task behavior in the classroom. This type of research is important because it allows direct comparisons between laboratory tests of attention and a criterion measure obtained in the setting where attention is often most problematic (e.g. school). The closer CPT data is to data obtained during direct observation of behavior in the natural setting, the more ecologically valid the results of the CPT are likely to be (Barkley, 1991). Thus, studies comparing the CPT to classroom observations would contribute valuable information regarding the validity of the CPT.

In one of the first studies of this kind, Kupietz and Richardson (1978), administered both an auditory vigilance task and a visual vigilance task to sixteen students between the ages of seven and twelve. A time sampling procedure was used to record the classroom behavior of the children during a small group remedial reading session.

Behaviors such as head or body turning, playing with objects, vocalizations and leaving the table or seat were recorded during two separate thirteen minute observations. Correlations were then computed between off-task behavior, commission errors, and omission errors. Results indicated that children who exhibited higher rates of off-task behavior during the reading session also made more vigilance errors. Interestingly, both omission and commission errors on the visual CPT correlated significantly with off-task behavior during reading (.56, $p<.05$ and .63, $p<.01$ respectively). Only omission errors on the auditory CPT correlated with off task behavior during the reading session (.66, $p<.01$). When chronological age was held constant, the partial correlations were slightly reduced for both the auditory and visual task and significance was only reported for the visual task (.46, $r<.05$ for omission errors and .49, $r<.05$ for commission errors). A correlation was also computed for the size of the vigilance decrement and observations of behavior. It was hypothesized that the size of the vigilance decrement would be positively correlated with off-task behavior. Results demonstrated that there was no significant

correlation between the size of the vigilance decrement (omission errors during the second half of the CPT minus omission errors during the first half of the CPT) and off-task classroom behavior.

Barkley (1991) also examined the relationship of the CPT to direct observational data. He conducted a post-hoc analysis of data collected during a previous study involving one hundred and forty children ages six to eleven. The group was comprised of children who had been classified as ADHD with hyperactivity, ADHD without hyperactivity, and non-ADHD controls. Each child completed a nine minute vigilance task (Gordon, 1983). The behavior of each child was observed and recorded during a fifteen minute period while the student sat alone in a lab playroom and completed a series of math tasks. Results indicated modest correlations (.12 to .34) between the behavior categories and scores on the CPT. Omission scores were significantly correlated with off-task (.32), fidgets (.26), and total ADHD categories (.34). Commission scores were significantly correlated with only the total ADHD behavior category (.32).

Barkley conducted another post hoc analysis that examined the relationship between CPT scores and direct observations of behavior (Fischer, Barkley, Edelbrock & Smallish, 1990). Data was collected from one hundred and sixty hyperactive and non-hyperactive subjects ages twelve to twenty years. Each subject completed a twelve minute vigilance task (Gordon 1987). The behavior of each subject was also observed and recorded during a fifteen minute session while the adolescent sat alone in a clinic room and completed a packet of math problems. Music was played during the session to serve as a distractor. Correlations between the CPT scores and the observation categories ranged from .03 to .44. Omission scores were not significantly correlated with any of the behavior categories. Commission scores were significantly correlated to off-task (.41), plays with objects (.26), out of seat (.44), and total ADHD behavior (.39) categories.

Finally, Prinz and his colleagues (1984) attempted to investigate the ecological validity of the CPT and the effects of distractions on CPT performance. Twenty nine ADDH boys, twelve boys classified as Learning Disabled and thirteen control boys were administered the CPT

(Rutschmann, Comblatt, & Erlenmeyer-Kimling, 1977). The children were then observed while completing a classroom analogue task (ANALOGUE) which was constructed to resemble independent seatwork. ANALOGUE consisted of a written discrimination task which was completed by the children while a color video monitor presented two alternating scenes. One scene was simply a blank screen while the other was a classroom activity scene (7 to 9 year old children engaging in classroom activities). The purpose of the fluctuating presentation on the monitor was to investigate the effects of distraction on the attention of ADDH children. For the entire sample of children, time on task and number correct were moderately and significantly correlated with CPT D-Prime ($p<.01$) for distraction and non-distraction conditions. For the ADDH group, none of the behavior observation categories were correlated with CPT performance in either of the two conditions. Prinz et al. concluded that the validity of his ANALOGUE task was comparable to that of the CPT.

While the studies above contribute significantly to the existing literature, it is important to recognize several limitations. First, and most importantly, only

Kupietz and Richardson collected observations of behavior in a true classroom setting with several classmates and a teacher present. The other researchers collected data in either a clinic testing room or laboratory playroom. Subjects were left alone during the observation period and did not experience the types of distractions typically encountered in a classroom (e.g. peer and teacher distractions, noise, etc). Thus, generalization of the findings is limited and a true comparison of performance on a CPT and behavior in the classroom was not conducted. Another concern is that the one study which examined student behavior in a real classroom setting did not examine behavior while the subjects were engaged in independent seatwork. Instead, Kupeitz and Richardson (1978) collected observation data while students were engaged in a remedial reading group. Off-task behavior (i.g. attention to task) was not directly coded. Instead, "head turning", "body turning", "leaving seat" and "playing with objects" was coded and used to infer off task behavior. It is possible that students could have been off task but not engaged in any of these behaviors and/or on task and simultaneously engaged in these behaviors. One

last concern is that subject selection may have confounded results in the Kupietz and Richardson study. For example, only sixteen students participated in the study. These subjects were referred because of reading problems and no information regarding attentional problems, academic problems or psychopathology was noted.

Unfortunately, the studies reviewed above do little to support the ecological validity of the CPT. To date, the literature lacks studies investigating the relationship between scores on continuous performance tests and observations of behavior obtained during independent seatwork in a non-simulated classroom observation.

ADHD vs Non-ADHD Control Groups. Studies comparing the performance of children classified as ADHD and children classified as non-ADHD are plentiful. These studies are relevant when discussing the validity of continuous performance tests as they provide another method by which to evaluate the validity of the instrument.

The type of CPT, as well as, the method by which CPT performance is interpreted varies greatly from study to study. The definition and classification criteria for Attention Deficit Hyperactivity Disorder varies across

studies as well. For example, some researchers examined the performance of "hyperactive" and normal children, while others compared the performance of "ADD+H, ADD-H, and non-ADD children". As a result of the diversity in classification criteria and dependent measures, it is difficult to compare findings and conclusions within the CPT literature.

Some researchers have interpreted CPT performance for groups of hyperactive and non-hyperactive youngsters by comparing the number of commission and/or omission errors made. Using this method of interpretation, several studies have demonstrated that hyperactive children and children classified as ADHD make more errors of omission on CPT tests relative to non-hyperactive and non-ADHD controls (Barkley, DuPaul, & McMurray, 1990; Fischer, Barkley, Edelbrock, & Smallish, 1990; Hooks et al, 1994; Horn, Wagner, & Ialongo, 1989; McLaren, 1990; O'Dougherty, Nuechterlein, & Drew, 1984; Sykes, Douglas, Weiss, & Minde, 1971; Sykes, Douglas, & Morgenstern, 1973). These results, however, remain equivocal as there have been several futile attempts to replicate the above findings (Koriath, Gualtieri, Van Bourgondien, Quade, & Werry, 1985; Shapiro,

Garfinkel, 1986; Schachar, Logan, Wachsmuth, & Chajezyk, 1988; Smith, Corkum, & Bryson, 1989; Werry, Elkind, & Reeves, 1987; Wherry, Paal, Jolly, Adam, Holloway, Everett, & Vaught, 1993).

Hyperactive children and children classified as ADHD have also been found to exhibit a higher number of commission errors on CPT tests than their non-hyperactive and non-ADHD peers (Barkley, DuPaul, & McMurray, 1990; Fischer et al, 1990; Hooks et al., 1994; Horn et al., 1989; O'Doughterty et al., 1984; Shapiro et al., 1986, and Sykes et al., 1971; and Sykes et al., 1987). Conflicting evidence for this finding also exists, however, as there have been several studies which have yielded contradictory results (McLaren et al, 1990; Schachar et al., 1988; Smith et al., 1989; Werry et al., 1987).

With regard to vigilance decrements, hyperactive children and children classified with ADHD have been shown to demonstrate a more rapid decline in performance over time on the CPT than non-hyperactive controls (Hooks, Milich & Lorch, 1994; Kupitz, 1976; Sykes et al, 1973). This is interesting as some researchers believe that change in CPT performance over time is a better measure of

sustained attention than is the overall number of errors made on a CPT (Halperin, 1991). Hooks et al, 1994 reported that there was a more rapid decline with regard to omission but not commission errors in a group of ADHD children. Hooks et al. (1994) also demonstrated that this difference was obvious only after the first block of CPT trials thus suggesting that the groups performed similarly at the beginning of the task. Sykes et al, 1973 reported evidence that the more rapid declines in performance demonstrated by hyperactive/ADHD children are due to a relative deficit in ability to detect target stimuli and not to changes in tendency to respond (Sykes et al., 1973). Unfortunately, the findings of the above reviewed studies are inconsistent with other researchers in the field, therefore, findings regarding vigilance decrements are, not unequivocal (Van der Meere and Sergeant, 1988; Schachar, Logan, Wachsmuth, & Chajczyk, 1988; O'Dougherty, Nuechterlein, & Drew, 1984).

Subjects classified as hyperactive have been found to make more anticipatory and multiple responses on both auditory and visual CPT tasks than non-hyperactive controls (Sykes et al, 1973). Hyperactive subjects have also been found to perform worse on the delay task of the Gordon

Diagnostic System (Gordon, 1979; McClure and Gordon, 1984) regardless of age, IQ, or experimental condition. This task yields information similar to commission errors in that it is said to measure impulsive responding (Gordon, 1979).

Differences within ADHD subgroups have also been noted. Barkley, DuPaul and McMurray (1990) examined the CPT performance of a group of children classified as attention deficit disorder with hyperactivity (ADD-H) and a group of children classified as attention deficit without hyperactivity (ADD+H). Results indicated that while the difference was not significant, the ADD+H children made almost twice as many commissions as the ADD-H children. The two groups did not differ on number of omission errors or on number of correct responses. Interestingly, although the ADD+H group performed significantly worse than normal controls on number of omissions and number of commissions, the ADD-H group did not perform significantly worse than the normal controls on any of the measures obtained.

In a similar study, Halperin, O'Brien, Newcorn, Healey, Pascualvaca, Wolf, & Young (1990) divided subjects into four groups: pure hyperactive (HYP), pure aggressive

(AGG), mixed hyperactive/aggressive (HYP/AGG) and normal controls. The subjects completed a twelve minute CPT task which required them to respond when they viewed the letter X preceded by the letter A. The HYP group made significantly more omission errors and demonstrated longer latency X-only commission errors (hit Xs that were not preceded by an A) than all other groups. This was interpreted as reflecting a larger degree of inattention (Halperin et al, 1988). The HYP/AGG group made significantly more A-not-X errors (hit letters other than X following an A) with short reaction times and more A-only errors with long reaction times. This was interpreted as reflecting a larger degree of impulsivity (See Halperin et al, 1988).

More recently, Nigg, Hindshaw, and Halperin (1996), examined the differences in the CPT performance of ADHD high aggressive, ADHD low aggressive, and normal control children. ADHD groups were found to differ from normal control subjects on inattention scores (X-only errors) and dyscontrol scores (errors other than X-only, A-only, or A-not-X only), but the subgroups differed significantly only on impulsivity scores (A-not-X only and A-only errors).

Overall, the literature provides some evidence for the validity of continuous performance tests as measured by differences between groups of hyperactive/inattentive children and normal controls. Additionally, there appears to be some evidence in support of differences in CPT performance within the ADHD group. However, conflicting results exist and comparisons across studies are clouded by varied dependent measures and varied group selection criteria. Corkum and Siegal summarized the literature on group differences and noted that the majority of studies reviewed reported a significant difference in the level of vigilance (defined by omission errors and d') between ADHD and non-ADHD controls. Additionally, they reported that there was "no clear evidence for the validation of a sustained attention deficit in ADHD children" as measured by a differential decline in performance as time on task progressed. Finally, Corkum and Siegal stated that task and situational variables were not controlled in the majority of the studies reviewed, thus, these results may have been confounded by these variables. Lossier, McGrath, and Klein (1996) also attempted to resolve the controversy over the frequency and type of CPT errors committed by children with

ADHD. They postulated that several factors have contributed to the conflicting literature. These factors include varied methodological standards, comorbid diagnoses, heterogeneity of comparison groups, small group sizes, and wide age ranges. After conducting a meta analysis of 26 studies conducted between 1973 and 1995, the authors concluded that children with ADHD made significantly more errors of omission and commission than normal children. When using signal detection theory parameters, they reported that children with ADHD were less sensitive to the difference between targets and non-targets than their normal counterparts, but demonstrated comparable response bias.

ADHD vs other clinical groups. Attentional problems are not unique to children with Attention Deficit Hyperactivity Disorder. Attention deficits are characteristic of various psychopathologies and disabilities. It is, thus, not surprising that children with learning disabilities, impulse-control problems, and various psychiatric disorders have all been found to perform poorly on continuous performance tests when compared to normal controls (Eliason & Richman, 1987; Klee

& Garfinkel, 1983). Overall, CPTs have been inconsistent in distinguishing between ADHD children and other clinical groups (Barkley, DuPaul & McMurray, 1990; Koriath, Gualtieri, Van Bourgondien, Quade, & Werry; 1985). For example, Barkley, DuPaul & McMurray (1990) reported that children classified as ADD+H performed significantly worse than children classified as learning disabled on number of commission and number of omissions, however, children classified as ADD-H performed significantly worse than learning disabled children on number of commissions only. This finding, as well as others like it, is important in that it has led some to question the diagnostic utility of continuous performance tests (Barkley, 1991).

Diagnostic Utility

Although the CPT was originally designed to assess vigilance deficits in adults with brain damage, has become one of the most popular clinic-based measures of sustained attention and is routinely used to diagnose Attention Deficit Hyperactivity Disorder (DuPaul et al, 1992). Unfortunately, the utility of continuous performance tests for diagnosing ADHD has not yet been adequately determined. Questions remain regarding the degree to which continuous

performance tests provide data that is consistent with other components of a diagnostic evaluation. Likewise, it is unclear whether continuous performance tests can be used to determine if an individual child may have Attention Deficit Hyperactivity Disorder.

In a preliminary analysis, Trommer, Hoeppner, Lorber, and Armstrong (1988) reported that within a group of 40 children who met the DSM-III criteria for ADD, 35% performed in the normal range. The researchers concluded that these children would have gone undetected if the CPT alone had been used for diagnostic purposes. Further research conducted by Trommer et al (1988) revealed that within a group of 14 children diagnosed with ADD, 5 performed within the "abnormal" range (at or below the fifth percentile), 5 performed within the "borderline" range (between the sixth and twenty-fifth percentile), and 4 performed within the "normal" range (above the twenty-fifth percentile) on the Vigilance Task of the Gordon Diagnostic System. Thus, 71% of the research group demonstrated CPT performance outside of the normal range. Trommer et al. (1988) also reported that within a group of 6 non-ADD subjects, 83% demonstrated CPT performance

outside of the normal range. Based on these findings, it was concluded that the CPT may yield both false positive and false negative results.

Gordon, DiNiro, and Mettelman (1988) found that the diagnostic hit rate of CPT commission error scores agreed with classifications based on parent and teacher ratings for 52% of a sample (N=74) of clinic-referred children. Children were diagnosed as ADHD on the basis of ratings at or above the 90th percentile on the Hyperactivity factor of the CBCL and/or the Inattention factor on the CBCL-TRF. This study elicits skepticism as to the diagnostic utility of continuous performance test as the obtained diagnostic hit rate was no better than chance.

More recently, DuPaul, Anastopoulos, Shelton, Guevremont, and Metevia (1992) attempted to determine the degree to which scores on a CPT and the MFFT agreed with parent and teacher ratings in a sample of children referred to an outpatient clinic. Children's scores on the Gordon Vigilance Task were classified as being in either the ADHD range (above the 93rd percentile) or the non-ADHD range. Percentage agreements were calculated. The percentage agreements for CPT scores and parent and teacher ratings

were 22% for CPT total correct score and 35% for CPT commission errors. Even when allowing the most liberal classification scheme which defined the child's performance being in the ADHD range on any of three clinic test scores (e.g. MFFT, CPT total correct, and CPT commission errors), nearly 40% of the children diagnosed as having ADHD were not correctly identified using the classification paradigm.

Finally, Fischer, Newby, and Gordon, (1995) compared children with Attention Deficit Hyperactivity Disorder with normal continuous performance test scores and ADHD children with abnormal scores. These researchers reported a 70% to 80% agreement rate between CPT classifications based on the Gordon Diagnostic System and ADHD diagnostic classifications based on clinical and rating scale criteria. The level of agreement varied depending on age range and specific age considered. The highest level of agreement occurred for children under twelve years of age. The researchers also reported that children with normal scores showed less inattention and more conduct and psychosomatic problems on the Conners Parent Rating Scale and Teacher Rating Scale ratings. Additionally, these children performed normally on other laboratory measures

and were less likely to respond positively to stimulant medication and less likely to do well on a higher dose than a lower dose when response was documented.

A discriminant function analysis is a statistical procedure that allows an investigation into whether it is possible to predict group membership on the basis of a variety of predictory variables (Hair, Anderson, & Tatham, 1997). This type of analysis also allows one to answer questions such as "what is the best combination of predictor variables to maximize difference among groups?". A discriminant function analysis using CPT scores as predictor variables would provide useful information regarding the diagnostic utility of CPTs (Hair et al. 1987). Surprisingly, only three studies have reported utilizing this statistical procedure to study CPT performance.

Koriath, Gualtieri, Van Bourgondien, Quade, and Werry (1984) utilized a discriminant function analysis in order to determine whether differences among subjects on dependent measures would discriminate among diagnostic groups. Groups were defined as ADHD with Hyperactivity, Conduct Disorder, Conduct Disorder with Hyperactivity,

Emotional Disorder and Emotional Disorder with Hyperactivity. The ADHD group was subdivided into pervasive hyperactivity and situation hyperactivity. Commission errors and correct response scores on the original version of the continuous performance test, the CPT (Rosveld et al., 1956), were included in the discriminant function analysis along with Actometer scores (Schulman and Reisman, 1959), Matching Familiar Figures Test scores (Kagan et al., 1964), ratings on the Conners Teacher Rating Scale (Conners, 1969), ratings on the Conners Parent Rating Scale (Conners, 1973), and results of the Routh Actvity Room (Routh et al., 1974). The researchers reported only that the results of the analysis were negative. No details regarding the type of discriminant function analysis were reported. Additionally, further details regarding the analysis were not reported.

The manual for the Test of Variables of Attention (TOVA; Greenberg, 1987) reports results of a discriminant analysis. The first included a total of seventy three youngsters with ADHD and normal control children matched on age and sex. Each child completed the T.O.V.A.. Parents and teachers of the children completed a ten item Conners

Parent-Teacher Questionnaire. The T.O.V.A. and the Conners Questionnaire together correctly classified 87% of normal an 90% of the AD-HD subjects with 13% false positives and 10% false negative. The T.O.V.A. alone classified 84% of the AD-HD subjects while the Conners Questionnaire correctly classified only 70%.

Another study, conducted by Greenberg (1989) involved children diagnosed as undifferentiated Attention Deficit Disorder (without hyperactivity; UADD). A total of twenty three children participated. A discriminant analysis including T.O.V.A. scores and results of the ten item Conners Parent-Teacher Questionnaire revealed that the T.O.V.A. correctly classified 83% of normals and 79% of UADD subjects. The T.O.V.A. alone reportedly classified only 65% of UADD subjects as compared to the Conners Parent-Teacher Questionnaire which correctly classified only 43%.

Summary

Although continuous performance tests have become one of the most widely used standardized measures for assessing Attention Deficit Hyperactivity Disorder (Meents, 1989), the literature regarding these clinic-based instruments is

fairly obscure. Results of studies are often conflicting and comparison across analyses is hindered by the use of varied versions of the instrument and multiple methods of interpreting performance.

While continuous performance tests have been repeatedly demonstrated to possess adequate reliability, the validity of these instruments remains controversial. The use of various methods of interpreting performance, the utilization of various criterion against which to compare the instruments, the inclusion of various subject populations, and the failure to consider the effects of extraneous factors that may have influenced performance (e.g. age, intelligence, presence of the experimenter etc.) have most likely resulted in the conflicting findings reported in the literature. In addition to concerns regarding validity, the diagnostic utility of continuous performance tests has also not been established. The ability of the instrument to discriminate between various populations is essential if it is to be considered valid for use as a diagnostic tool.

METHOD

Participants

Second, third and fourth grade teachers from several local schools in East Baton Rouge, Jefferson, Assumption and Livingston parishes in southern Louisiana were asked to nominate potential participants for this study. Teachers were asked to nominate children who were exhibiting significant difficulties with attention and\or hyperactivity, as well as, children who were considered to be "average" students. From an initial sample of approximately 170 nominations, written consent was obtained for 120 children and their parents to participate in this study. All participants met the following criteria: 1)were between the ages of seven and nine, 2) possessed no evidence of auditory, language, or visual difficulties, 3) had never been diagnosed with a Pervasive Developmental Disorder, Developmental Language Disorder, or Cerebral Palsy, 4) had never been diagnosed with mental retardation, 5) had never been retained for academic reasons 6) were not currently failing any academic subjects. 7) were not presently receiving special education services. The total number of participants consisted of 86 children and

included 60 males (69.8%) and 26 females (30.2%). Two students were in the first grade, forty two students were in the second grade, twenty three students were in the third grade and eighteen students were in the fourth grade.

Several of the participants had previously been prescribed stimulant medication. The behavioral effects of pemoline and d-amphetamine last approximately six and ten hours respectively (Pelham, 1993). Children who were taking these or other similar medications were excluded from the study. Children who were regularly taking methylphenidate were allowed to participate as the behavioral effects of this medication disappear approximately four hours following ingestion. Medication effects are not likely if medication is discontinued at least twelve hours prior to testing. Thus, children who were presently taking stimulant medication were asked to refrain from taking medication on the days when data was collected in order to control for medication effects.

Measures

Conners Continuous Performance Test. The Conners Continuous Performance Test (CPT) is visually presented on a computer monitor and lasts approximately fourteen minutes.

Various letters of the alphabet appear one at a time and subjects are asked to "quickly press the space bar" for any letter except the letter X. The parameters of the task can be varied by the examiner, however, the standard task consists of six blocks with three sub-blocks of twenty stimuli each. The interstimulus interval varies across blocks and is either two or four seconds. The display time for each letter remains constant at 250 milliseconds. For the purpose of this study, the standard task of the CPT was administered according to the guidelines in the manual. The CPT allows interpretation of several variables including hits (number of appropriate responses to target stimuli), omissions (number of times subject fails to respond to target), commissions (number of responses to non-target), hit rate (mean response time for all target responses over all six time blocks), variability of SEs(standard deviation of the 18 standard error values calculated for each sub-block; another measure of response time), attentiveness (d'a measure of how well one discriminated between targets and non-targets), and risk taking(B a measure of response tendency). Results are presented using raw scores, T-scores and percentiles.

ADHD Rating Scale IV. The ADHD Rating Scale IV (DuPaul et al., 1994) consists of eighteen items based on DSM-IV criteria for Attention Deficit/Hyperactivity Disorder. The measure asks parents and/or teachers to rate the frequency of each item as occurring *not at all*, *sometimes*, *pretty much*, or *very much*. Normative data is available for age and gender. Information can be interpreted using the Inattention scale or the Hyperactivity scale. A total score is also available. Both scales were found to have adequate internal consistency (all coefficient alphas > .80) and test-retest reliability (DuPaul, 1996). Scores on both scales have been noted to correlate significantly with ratings on the Conners Parent and Teacher Rating Scales, as well as with direct observations of classroom attention and productivity. Preliminary factor analyses support a two-factor model for both scales that conforms with the breakdown of symptoms in the DSM-IV (DuPaul, 1996).

Conners' Parent Rating Scale-Revised (L). The Conners' Parent Rating Scale-Revised (CTRS; Conners, 1997) consists of 80 items and is often used to assess the frequency of various childhood behaviors. Parents of

children are asked to rate behaviors as being exhibited never, seldom, occasionally or often or very often by their child. Normative data is available for both age and sex.

Conners' Teacher Rating Scale-Revised (L). The Conners' Teacher Rating Scale-Revised (CTRS-R; Conners, 1997) consists of 59 items and is similar to the Conners' Parent Rating Scale-Revised. Teachers are asked to rate various student behaviors as being exhibited "never", "occasionally", "often" or "very often" by a particular student in the classroom. Norms are available for both age and sex.

Behavior Evaluation Tool. The Behavior Evaluation Tool (BET; de Back, 1997) is a rating scale which asks parents (BET-P) and/or teachers (BET-T) to identify recent behaviors exhibited by a child and to assess the severity of impairment for each behavior. The parent and teacher versions of the BET consist of 104 items each. Most of the items were directly adapted from the DSM-IV criteria lists for one of the following disorders: Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder, Separation Anxiety Disorder, Social Phobia, Generalized Anxiety Disorder, Major Depressive Disorder and Dysthymic

Disorder. Informants are asked to rate the severity of impairment ("size of the problem") for each behavior. When assessing "size of the problem" raters are asked to consider both the frequency of each behavior, as well as, the overall effect of each behavior on the child's ability to communicate and cooperate with family, teachers, friends, and classmates in the home, school, and at work. Items are rated using a five point likert-type scale ranging from "no problem exists" to "major problem for communication and cooperation in almost all or all areas of interaction on a consistent and frequent basis". Psychometric data on the BET is presently not available. The BET is included as it provides a quick rating of all of the most common childhood disorders based on DSM criteria and in order to begin collecting preliminary psychometric data.

Curriculum Based Measurement. Brief curriculum based measurement probes (Shinn, 1989) in the areas of math and reading were conducted individually with each participant prior to the classroom observation. The probes provided information regarding reading and math fluency and accuracy. Information gathered during the math probes was

used to prepare instructional level math sheets which were completed by participants during a classroom observation.

During the reading probe, each student was asked to read aloud for one minute from his or her basal reader. The percentage of words read correctly during the one minute interval was then calculated and used as a dependent measure. During the math probe, participants were asked to complete worksheets consisting of math problems selected from grade level text for two minutes. The number of problems correctly completed within the two minute time period was recorded and used as a dependent measure. Students who were found to be functioning below instructional level in the area of math were administered additional math probes until their instructional level was determined. Packets containing instructional level multi-skill math problems were constructed and completed by the participant during the classroom observation. Instructional level was defined as the level at which the student was able to complete between 70% and 85% of the problems correctly (Gickling and Rosenfeld, 1995).

Classroom Observations. Observers were selected and trained in the use of an observational coding system which

has been frequently used to assess the classroom behavior of children with Attention Deficit Hyperactivity Disorder (Barkley, 1991). Training consisted of instruction and repeated practice observations under the supervision of a trainer. When observers achieved 80% reliability with the trainer on practice observations they were allowed to conduct classroom observations of participants.

One ten minute classroom observation was obtained for each participating child. Behaviors were observed and recorded using a ten second partial interval time sampling procedure. Target behaviors and response definitions were based on the ADHD observational code described by Barkley (1991). Specifically, behaviors that were observed included off-task, fidgeting, out of seat, playing with objects, and inappropriate vocalizations. In addition, peer and teacher attention was recorded as a procedural integrity measure. Behavioral definitions can be found in Appendix B.

All observations were conducted within two weeks of the CPT administration. The observations occurred at the time of day corresponding to that during which the continuous performance tests were completed (e.g. morning or afternoon). Observations were conducted during a period

of time when the target student and the remainder of the class was engaged in independent seatwork.

During each observation, the participant was instructed by his or her teacher to complete a packet of prepared multi-skill math sheets. The difficulty level of the math sheets were consistent with the instructional level of participant. Instructional level was determined using curriculum based assessment probes as described above.

In order to measure the reliability of the observations, two independent observers simultaneously collected data on a random sample of 21% of all classroom observations. Reliability was calculated on an interval by interval basis for each of the behavior categories by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying this number by 100 (Cooper, 1987). Reliability estimates ranged from to 94.2% to 100%. The reliability estimates averaged 94.2% for off task behavior, 99.6% for inappropriate vocalizations, 100% for out of seat, 99.8% for playing with objects, and 99.8% for fidgeting These agreement estimates

are very acceptable and are indicative of high reliability between the raters.

Procedure

Parent and Teacher Interviews. Parents of potential participants were contacted by letter or telephone and the purpose and details of the study were explained. Those agreeing to participate were asked to complete a consent form (Appendix A). They were also asked to complete the Conners' Parent Rating Scale-Revised, ADHD Rating Scale IV, and the BET-P scale. Teachers of the participating students were also provided with an explanation of the study and asked to complete a packet of rating scales which included the Conners' Teacher Rating Scale-Revised (CTRS-R), the ADHD Rating Scale IV, and the BET-T scale. The CTRS-R, CPRS-R, and ADHD Rating Scale IV were used to determine eligibility for the study and was used as dependent measures. Parents and teachers of children who were taking medication were asked to rate the child's behavior as observed when he or she was without medication.

Group Classification. Data gathered from the CPRS-R, CTRS-R, the BET and the ADHD Rating Scale IV were used to assign each of the participants to either a No

Diagnosis/Control group or an Attention Deficit Hyperactivity Disorder (ADHD) group. Children were assigned to the No Diagnosis/Control group if the following criteria were met: (a) absence of a history of mental health services for behavioral difficulties as reported by parents, (b) teacher and parent endorsements of *pretty much* or *very much* on less than six of the nine items which assess inattention on the ADHD Rating Scale IV, and c) teacher and parent endorsement of less than six of the nine items which assess hyperactivity-impulsivity on the ADHD Rating Scale IV d) A T-score below 65 on all scales of the Conners' Teacher Rating Scale-Revised and the Conners' Parent Rating Scale-Revised.

Children were included in the Attention/Deficit Hyperactivity Disorder group if they exhibited a significant number of ADHD symptoms in the school and home setting. Children were determined to be exhibiting a significant number of ADHD symptoms in the school setting if at least one of the following criteria were met (a) teacher endorsements of *pretty much* or *very much* on six or more of the nine items which assess inattention on the ADHD Rating Scale IV, or (b) teacher endorsements of *pretty much*

or *very much* on six or more of the nine items which assess hyperactivity-impulsivity on the ADHD Rating Scale IV, or c) a T score greater than 65 on either the Conners' ADHD Index, the Cognitive Problems Index, the DSM-IV Inattentive Index, the DSM-IV Hyperactive-Impulsive Index or the Hyperactivity Index of the Conner's Teacher Rating Scale-Revised, d) teacher endorsement of pretty much or very much on six or more of the items listed on the BET-T which were derived from the DSM-IV criteria necessary for a diagnosis of Attention Deficit Hyperactivity Disorder Predominately Inattentive Type or Attention Deficit Hyperactivity Disorder Predominately Hyperactive-Impulsive Type.

Children were determined to be exhibiting a significant number of ADHD symptoms in the home setting if at least one of the following criteria were met: (a) parent endorsements of *pretty much* or *very much* on six or more of the nine items which assess inattention on the ADHD Rating Scale IV, or (b) parent endorsements of *pretty much* or *very much* on six or more of the nine items which assess hyperactivity-impulsivity on the ADHD Rating Scale IV, or c) a T score greater than 65 on either the ADHD Index, the Cognitive Problems Index, the DSM-IV Inattentive Index, the

DSM-IV Hyperactive-Impulsive Index or the Hyperactivity Index of the Conner's Parent Rating Scale-Revised, d) Parent endorsements of *pretty much* or *very much* on six or more of of the items listed on the BET-P which were derived from the DSM-IV criteria necessary for a diagnosis of Attention Deficit Hyperactivity Disorder Predominately Inattentive Type, Attention Deficit Hyperactivity Disorder Predominately Hyperactive-Impulsive Type. In addition to the criteria listed above, children included in the Attention Deficit Hyperactivity Disorder group were required to meet the following criteria: (a)problem behaviors endorsed on the ADHD Rating Scale IV must have been exhibited for at least six months according to parent and/or teacher report and (b) onset of the problem behaviors endorsed on the ADHD Rating Scale IV must have been before the age of seven.

Twenty three of the 86 participants met the classification criteria for the No Diagnosis/Control group while 45 children met the classification criteria for the Attention Deficit/Hyperactivity Disorder group (i.e. difficulties in both the home and school setting). Interestingly, 18 children exhibited a significant number

of ADHD symptoms in the school setting only. For some of the statistical analyses, children in this "school difficulties" group were included with children in the ADHD group.

ADHD Subgroups

In order to conduct statistical analyses with the subtypes of Attention Deficit Hyperactivity Disorder listed in the DSM-IV, the ADHD group, was further subdivided into three groups, Attention Deficit Hyperactivity Disorder Predominantly Inattentive Type, Attention Deficit Hyperactivity Disorder Predominantly Hyperactivity-Impulsivity Type and Attention Deficit Hyperactivity Disorder Combined Type using the ADHD Rating Scale IV. Subjects were assigned to an Attention Deficit/Hyperactivity Disorder Predominately Inattentive Type (ADHD-I) group if their teacher rated six or more of the items on the ADHD Rating Scale IV inattention factor as occurring "pretty much" or "very much" and less than six of the items on the ADHD Rating Scale IV hyperactivity-impulsivity factor as occurring "pretty much" or "very much". Subjects were assigned to an Attention Deficit/Hyperactivity Disorder Predominately

Hyperactive-Impulsive Type (ADHD-HI) if their teacher rated six or more of the items on the ADHD Rating Scale IV hyperactive-impulsive factor as occurring "pretty much" or "very much" and less than six of the items on the ADHD Rating Scale IV inattentive factor as occurring "pretty much" or "very much". Finally, subjects were assigned to an Attention Deficit/Hyperactivity Disorder Combined Type (ADHD-COM) if six or more items on the inattention factor and six or more of the items on the hyperactivity-impulsivity factor were rated as occurring "pretty much" or "very much". When teachers did not rate six or more items as occurring "pretty much" or "very much" on either the inattention factor or the hyperactivity-impulsive factor of the ADHD Rating Scale IV, rating scale data from the Conners Teacher Rating Scale-Revised was used to classify subjects into their respective groups.

Medication

Participants who regularly received methylphenidate or Ritalin were asked to consult their prescribing physician in order to obtain permission to discontinue medication for twelve hours prior to the administration of the CPT and classroom observation. This was to ensure the

absence of medication effects during data collection while only minimally disrupting the participants regular medication regime.

Continuous Performance Test Administration

All participants were asked to attend a testing session either at a local psychology clinic or at their school. At that time, the children individually completed the standard version of the Continuous Performance Test (CPT; Conners, 1995). The Conners' CPT standard task last approximately 14 minutes and consists of an 1.5 second ISI, a 200 millisecond display time, and a target probability of 10%. The test was administered according to the guidelines printed in the Conners' CPT Computer Program 3.0 User's Manual (Conners, 1995). Participants completed the standard practice test prior to the administration of the test (Conners, 1995). Participants obtaining an omission score greater than 70% on the standard practice test were provided with repeated instructions as it is stated that this score indicates that the instructions were misunderstood (Conners, 1995). Examiners remained in the testing room during the standard practice test and the test. Examiners stood behind the

participant during the administration of the CPT. The number of omission errors, commission errors and mean response times was recorded upon completion of the test. Other measures that were obtained included attentiveness (d') and risk taking (B).

RESULTS

Correlation Analyses

No Diagnosis/Control and ADHD Group. In order to examine the relationship between CPT performance, systematic observations of classroom behavior, and data obtained during CBM probes, two simple correlation analyses were conducted. The first analysis was conducted using data obtained from children in the No/Diagnosis Control group (N=23) and the ADHD group (N=45). The following variables were included in the correlation analyses: CPT omission scores, CPT commission scores, CPT total hits, CPT variability of SE, number of intervals off task, number of intervals engaged in inappropriate vocalizations, number of intervals playing with objects, number of intervals fidgeting, number of intervals out of seat, number of intervals engaged in total ADHD behavior (i.e. off task, out of seat, inappropriate vocalizations and playing with objects), number of intervals engaged in disruptive behavior (e.g. out of seat, vocalizing, or playing with objects), number of digits completed correctly during the observation, number of digits completed correctly on grade level multi-skill math probes

During one minute, number of words read correctly during one minute on grade level reading text, and accuracy of reading during one minute on grade level reading text.

A summary of all correlations for these variables appears in Table 1. Significant correlations were found between CPT performance and several classroom behaviors. Specifically, CPT omission scores were positively correlated with off task behavior (.317; p=.008), inappropriate vocalizations (.455; p=000), out of seat behavior (.332; p=006), playing with objects (.254; p=.036), total ADHD behaviors (.385; p=.001), and disruptive behaviors (.548; p=.000). Because the number of total hits on the CPT is directly related to the number of ommissions, correlations between total hits and observation data were the same as correlations between omission scores and observation data but were negative. CPT Commission scores were not found to be significantly correlated with any of the variables examined. CPT variability SE scores were correlated with off task behavior (.265;p=.029), inappropriate vocalizations(.359;p=.001), out of seat (.251;p=.039), total ADHD behaviors (.326;p=.007), and in disruptive

Table 1

Correlation Matrix of CPT scores, Observation data, and CBM scores

Variable	1	2	3	4	5	6	7	8
CPT								
1.Omission		.01	-1.0**	.86**	.32**	.46**	.33**	-.08
2.Commission			-.01	.09	.03	.06	-.05	.01
3.Total				-.87**	-.32**	-.46**	-.33**	.09
4.Var SE					.27*	.40**	.25*	-.10
Observation								
5.Off						.40**	-.01	.47**
6.Voc							.36**	.01
7.Out Seat								-.08
8.Fidget								1.0

(table cond.)

Note. * p< .05
** p< .01

Variable	9	10	11	12	13	14	15
CPT							
1.Omission	.25*	.39**	.55**	-.21	-.28*	-.15	-.34**
2.Commission	.03	-.00	-.01	-.16	-.04	-.17	-.05
3.Total	-.25*	-.39**	-.55**	.21	.28*	.15	.34**
4.Var SE	.16	.33**	.46**	-.08	-.34**	-.07	-.37**
Observation							
5.Off	.51**	.79**	.29*	-.32**	.02	-.25*	-.12
6.Voc	.33**	.52**	.74**	-.11	-.09	-.12	-.21
7.Out Seat	.11	.32**	.74**	-.21	-.08	-.07	-.14
8.Fidget	.36**	.66**	.04	-.08	.15	-.13	.04
9.Objects		.49**	.41**	-.10	.03	-.04	.03
10.Tot Bx			.63**	-.30*	.03	-.19	-.16
11.Disr Bx				-.22	-.11	-.07	-.16
12.Dig Cor					.12	.41*	.23
CBM							
13. Read Acc						.05	.60**
14. Math Dig							.17
15. WCPM							1.0

Note. * p< .05
** p< .01

behaviors (.456;p=.000). CPT variability SE scores are a measure of response time consistency.

Significant correlations were also found between CPT scores and CBM data. CPT omission scores were negatively correlated with reading accuracy (-.278;p=.030) and with words correct per minute (-.344;p=.005) during one minute grade level reading probes. CPT variability SE scores were also negatively correlated with reading accuracy (-.335;p=.008) and words correct per minute (-.366;p=.003) during the reading probes. Total hits were significantly and positively correlated with reading accuracy and words correct per minute. Correlations were (.278; p=030) and (.344; p=.005) respectively.

No Diagnosis/Control Group, ADHD Group and School Difficulties Group. The second correlation analysis was conducted using data gathered from children in the ADHD group, the School Difficulties group, and the No Diagnosis/Control group (N=86). The same variables that were included in the first set of analyses were again used. Results of the second analysis were similar to those of the first and can be found in Table 2. CPT data were again found to be significantly correlated with classroom

Table 2

Correlation Matrix of CPT scores, Observation data, and CBM scores

Variable	1	2	3	4	5	6	7	8
CPT								
1.Omission		.03	-1.0**	.84**	.31**	.38**	.33**	-.06
2.Commission			-.03	.09	-.06	.04	-.01	.02
3.Total				-.84**	-.31**	-.38**	-.33**	.06
4.Var SE					.29**	.41**	.25*	-.10
Observation								
5.Off						.44**	.01	.37**
6.Voc							.31**	.02
7.Out Seat								-.07
8.Fidget								1.00

(table cond.)

Note. * p< .05
** p< .01

Variable	9	10	11	12	13	14	15
CPT							
1.Omission	.26*	.37**	.53**	-.22*	-.25*	-.16	-.34**
2.Commission	.06	-.05	-.03	-.18	-.04	-.14	-.07
3.Total	-.26*	-.38**	-.53**	.22*	.25*	.16	.34**
4.Var SE	.19	.33**	.46**	-.14	-.28*	-.08	-.33**
Observation							
5.Off	.54**	.78**	.33**	-.32**	.03	-.24	-.11
6.Voc	.40**	.51**	.73**	-.12	-.05	-.08	-.14
7.Out Seat	.11	.36**	.72**	-.20	-.07	-.06	-.12
8.Fidget	.31**	.64**	.04	-.10	.13	-.10	.02
9.Objects		.51**	.45**	-.12	.03	-.05	.02
10.Tot Bx			.64**	-.31*	.01	-.16	-.13
11.Disr Bx				-.22	-.08	-.05	-.13
12.Dig Cor					.22*	.41*	.22*
CBM							
13. Read Acc						.27*	.63*
14. Math Dig							.30**
15. WCPM							1.0

Note. * p< .05
** p< .01

observation data. Specifically, CPT omission scores were positively correlated with off task behavior (.305;.p=004), inappropriate vocalizations (.384;.p=000), playing with objects(.262;p=.015), out of seat behavior (.330;p=.002), total ADHD behavior (.374;p=.000), and disruptive behavior (.527;p=.000). CPT variability SE scores were significantly correlated with number of intervals off task (.285;.p=008), number of intervals engaged in inappropriate vocalizations (.407;.p=000), number of intervals out of seat (.245;p=.023), number of intervals engaged in total ADHD behavior (.326;p=.002), and number of intervals engaged in disruptive behavior (.464;p=.000).

CPT data were also significantly correlated with CBM data. Specifically, negative correlations were found between reading accuracy during a one minute reading probe and CPT omission scores(-.250;p=.034) and CPT variability SE scores (-.279;p=.017). Total hits on the CPT was positively correlated with reading accuracy (.250;p=.034). Words correct per minute was significantly correlated with CPT omission scores (-.341;p=.002) and CPT total hits (.341;p=.002).

Behavior Evaluation Tool. A third set of correlation analyses was conducted with the ADHD group and the No Diagnosis/Control group in order to gather preliminary psychometric data on the BET rating scale. Specifically, correlations were computed to investigate the relationship between several sections of the BET teacher rating scale and the BET parent rating scale with various factors on the Conners' Teacher Rating Scale-Revised, the Conners' Parent Rating Scale-Revised, and the parent and teacher forms of the ADHD-Rating Scale IV. Results of all correlations can be found in Table 3. Section A (ADHD) of the BET-T (teacher rating scale) was significantly correlated with the Inattentive factor (.951; p=.000), Hyperactive-Impulsive factor (.948; p=.000) and the total score (.927; p=.000) of the ADHD Rating Scale IV teacher form. Significant positive correlations were also found between Section A of the BET-T and the Cognitive Problems/Inattention factor (.794; p=.000), the Hyperactivity factor (.834; p=.000), the Conners' ADHD Index factor (.910; p=.000), the DSM-IV Inattentive factor (.870; p=.000), and the DSM-IV Hyperactive-Impulsive

Table 3

Correlation Matrix of BET-T, ADHD Rating Scale IV (teacher version), and Conners' Rating Scale Revised (teacher form).

Variable	1	2	3	4	5	6	7
BET-T							
1.Section A (ADHD)		.65**	.45**	.42**	.95**	.95**	.93**
2.Section B (Opp)			.47**	.59**	.57**	.68**	.61**
3.Section C (Anxiety)				.73**	.41**	.35**	.36**
4.Section D (Depression)					.37**	.36**	.41**
ADHD-IV (T)							
5.Inattentive						.89**	.93**
6.Hyp-Imp							.92**
7.Total Score							

(table cond.)

Note. * p< .05
** p< .01

Variable	9	10	11	12	13	14	15	16
BET-T								
1.Section A (ADHD)	.60**	.79**	.83**	.41**	.91**	.87**	.84**	.89**
2.Section B (Oppositional)	.88**	.60**	.74**	.50**	.73**	.64**	.76**	.71**
3.Section C (Anxiety)	.26*	.40**	.32*	.63**	.38**	.40**	.31*	.36**
4.Section D (Depression)	.39**	.35**	.40**	.54**	.42**	.35**	.42**	.39**
ADHD-IV (T)								
5.Inattentive	.55**	.81**	.80**	.42**	.89**	.90**	.81**	.89**
6.Hyp-Imp	.68**	.75**	.87**	.38**	.91**	.83**	.89**	.89**
7.Total Score	.56**	.79**	.79**	.41**	.85**	.83**	.80**	.84**
CTRS-R								
9.Oppositional		.55**	.75**	.51**	.73**	.64**	.78**	.72**
10.Cog Problems			.71**	.41**	.81**	.89**	.72**	.83**
11.Hyperactivity				.47**	.94**	.83**	.98**	.93**
12.Anxious-Shy					.43**	.44**	.45**	.45**
13.ADHD Index						.97**	.95**	.97**
14.DSM-IV Inatt							.85**	.95**
15.DSM-IV Hyp-Imp								.95*
16.DSM-IV Tot								

Note. * p< .05;
** p< .01

factor of the Conners' Teacher Rating Scale-R (.887; p=.000). Section B (Oppositional) of the BET-T was found to be significantly correlated with the Oppositional factor of the CTRS-R (.877; p=000) and Section C (Anxiety) of the BET-T was found to be significantly correlated with the Anxiety factor (.634; p=.000) of the CTRS-R. There were no comparable factors for Section D (Depression) of the BET-T.

Significant correlations between the parent version of the BET (BET-P) and factors on other parent rating scales were also found (Table 4). For example, Section A of the BET-P was found to be significantly correlated with the Inattention factor (.923; p=.000), the Hyperactive-Impulsive factor (.888; p=.000) and the total score (.944; p=000) on the ADHD Rating Scale IV parent form. This section of the BET-P was also significantly correlated with the Cognitive Problems/Inattention factor (.737; p=.000), the Hyperactivity factor (.838; p=.000), the Conners' ADHD Index factor (.821; p=.000), the DSM-IV Inattentive factor (.817; p=.000), and the DSM-IV Hyperactive-Impulsive factor of the Conners' Teacher Parent Scale-R (.804; p=.000). Section B (Oppositional)

Table 4

Correlation Matrix of BET-P, ADHD Rating Scale IV (parent version) and the Conners' Rating Scale Revised (parent form).

Variable	1	2	3	4	5	6	7
BET-P							
1.Section A (ADHD)		.87**	.57**	.44**	.92**	.89**	.94**
2.Section B (Opp)			.54**	.54**	.79**	.84**	.85**
3.Section C (Anxiety)				.71**	.47**	.43**	.47**
4.Section D (Depression)					.40**	.41**	.42**
ADHD-IV (P)							
5.Inattentive						.83**	.96**
6.Hyp-Imp							.96**
7.Total Score							

(table cond.)

Note. * p< .05
** p< .01

Variable	9	10	11	12	13	14	15	16
BET-P								
1.Section A (ADHD)	.78**	.74**	.84**	.40**	.82**	.82**	.80**	.80**
2.Section B (Oppositional)	.84**	.62**	.78**	.40**	.71**	.72**	.75**	.78**
3.Section C (Anxiety)	.60*	.33**	.54*	.64**	.49**	.51**	.55*	.56**
4.Section D (Depression)	.53**	.20	.43**	.48**	.30**	.34**	.38**	.40**
ADHD-IV (P)								
5.Inattentive	.70**	.81**	.74**	.39**	.81**	.86**	.74**	.84**
6.Hyp-Imp	.78**	.64**	.89**	.32**	.76**	.73**	.87**	.84**
7.Total Score	.78**	.76**	.85**	.37**	.82**	.85**	.85**	.89**
CPRS-R								
9.Oppositional		.62**	.81**	.55**	.74**	.71**	.84**	.82**
10.Cog Problems			.62**	.41**	.86**	.92**	.64**	.83**
11.Hyperactivity				.38**	.79**	.72**	.95*	.89**
12.Anxious-Shy					.52**	.52**	.40**	.48**
13.ADHD Index						.92**	.85**	.95**
14.DSM-IV Inatt							.75**	.93*
15.DSM-IV Hyp-Imp								.93*
16.DSM-IV Tot								

Note. * p< .05;
** p< .01

of the BET-P was found to be significantly correlated with the Oppositional factor of the CPRS-R (.841; p=000) and Section C (Anxiety) of the BET-P was found to be significantly correlated with the Anxiety factor of the CTRS-IV (.638; p=.000). There were no comparable factors for Section D (Depression) of the BET-P.

While significant and impressive correlations were found when various sections of the BET-P and the BET-T were compared with similar factors of other rating scales, it is important to note that significant correlations were also found when various sections of the BET-P and BET-T were compared with dissimilar factors of other rating scales. For example, Section C (Depression) of the BET-T was significantly and positively correlated with the Oppositional factor of the CTRS-R (.256; p=.041).

Comparison of Means.

ADHD and No Diagnosis/Control. In order to determine whether there were significant differences between diagnostic group, and age and grade, Analyses of Variance were performed on the data. ANOVAs included children in the ADHD group (N=45) and the No Diagnosis/Control group (N=23). Results indicated that there were no significant

differences between the group means with respect to age [F(1,66)=.039, p>.05] and grade [F(1,66)=.193,p>.05]. In order to determine if there were significant differences between diagnostic group, and gender and household income level, Chi-square analyses were performed. Results of these analysis suggested that there were no significant difference between the groups with regard to the variable gender $X^2=(1)=.759,p<.05$) or household income $X^2=(6)=.339,p<.05$.

ADHD plus School Difficulties and No Diagnosis/Control. Another set of analyses was conducted using the ADHD group plus with the School Difficulties group (N=63) and the No Diagnosis/Control group (N=23). ANOVAs were conducted using the variables age and grade to investigate the possibility of significant differences between the group means. Results of these analyses suggested no significant differences between the groups with respect to age ([F(1,84)=.001, p>.05] and grade [F(1,84)=.464, p>.05]. Chi-square analyses were conducted using the variables group, and household income, and gender. Results indicated no significant differences

between the groups with regard to the variables gender $X^2=(1)=.579, p<.05)$ and household income $X^2=(6)=.540, p<.05)$

Discriminant Function Analyses

Several standard discriminate function analyses were completed in an attempt to answer the following research questions: a) How well do scores on the CPT predict membership into diagnostic groups?, b) How well do classroom observations predict membership into diagnostic groups?, c) How well does data obtained from CBM probes predict membership into diagnostic groups?, d) What is the best combination of variables (CPT scores, classroom observation data, and CBM data) for predicting membership into diagnostic groups?

ADHD Combined Type group and No Diagnosis/Control. A direct discriminate function was performed using the No Diagnosis/Control group (N=23) and the children in the ADHD Combined Type group (N=28). Because of the relatively small number of children in the sample who were assigned to the Attention Deficit/ Hyperactivity Disorder Predominately Hyperactive-Impulsive Type group (ADHD-HI; N=9), and because recent literature suggests that the validity of this subtype is neither empirically supported

nor clinically useful (Power & DuPaul, 1996), this subtype was not included in the analyses. The Attention Deficit Hyperactivity Disorder Predominately Inattentive group was also very small (N=8) and was not included in this discriminant function analysis.

The number of omission scores on the CPT was used as the predictor of group status. Omission scores correctly classified 64.7% of the subjects according to initial group membership. This percentage is only slightly better than that expected by chance (50%). The percentage of subjects classified according to their initial group membership by CPT omission scores is presented in Table 5. Results suggest that the omission scores accurately classified 60.7% of the ADHD combined and 69.6% of the ADHD control subjects.

Another discriminant function was conducted using CPT commission scores as a predictor variable. The percentage of subjects classified according to their initial group membership by commission scores is presented in Table 6. Only 62.7% of the subjects were correctly classified according to their initial group membership by teacher rating scale data. This was somewhat expected given the

Table 5

Cases Predicted by Omission scores on the Conners' Continuous Performance Test.

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (combined type)	28	17 (60.7%)	11 (39.3%)
Group 2 Control	23	7 (30.4%)	16 (69.6%)
Percent of cases correctly classified:		64.7%	

Table 6

Cases Predicted by Commission scores on the Conners' Continuous Performance Test

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (Combined type)	28	22 (78.6%)	6 (21.4%)
Group 2 Controls	23	13 (56.5%)	10 (43.5%)
Percent of cases correctly classified:		62.7%	

absence of a significant difference between the group means for commission scores. A discriminant function analyses using the variable, CPT Variability SE scores as a predictor resulted in a somewhat higher classification accuracy. CPT variability SE scores correctly classified 68.6% of the sample according to their initial group membership by teacher rating scale data (Table 7).

Discriminant function analyses was conducted using the number of intervals during which subjects were engaged in off task behavior during the classroom observation as the predictor variable. The percentage of subjects correctly classified according to their initial group membership was 62.7% and was again not much better than chance (Table 8).

Finally, two separate discriminant function analyses using the number of words read correctly per minute on grade level reading probes and number of digits completed correctly per minute during one minute grade level math probes were conducted. Results of these analyses can be found in Tables 9 and 10. These analyses resulted in classification accuracies of 63.3% and 55.1% percent respectively.

Table 7

Cases Predicted by Variability SE scores on the Conners' Continuous Performance Test

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (Combined type)	28	15 (53.6%)	13 (46.4%)
Group 2 Controls	23	3 (13%)	20 (87%)
Percent of cases correctly classified:		68.6%	

Table 8

Cases Predicted by off task behavior during classroom observation.

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (Combined type)	28	24 (85.7%)	4 (14.3%)
Group 2 Controls	23	15 (65.2%)	8 (34.8%)

Percent of cases correctly classified: 62.7%

Table 9

Cases Predicted by words correct per minute on grade level CBM probes

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (Combined type)	26	17 (65.4%)	9 (34.6%)
Group 2 Controls	23	9 (39.1%)	14 (60.9%)

Percent of cases correctly classified: 63.3%

Table 10

Cases Predicted by math digits correct on grade level CBM probes

Actual Group	# of cases	Predicted Group Membership 1	2
Group 1 ADHD (Combined type)	27	23 (85.2%)	4 (14.8%)
Group 2 Controls	22	18 (81.8%)	4 (18.2%)

Percent of cases correctly classified: 55.1%

The best predictor of all variables studied was CPT variability SE scores. The prediction accuracy of the CBM probes, CPT omission scores and off task behavior were all similar. Overall, none of the variables examined appear to be very good predictors of group membership according to the discriminant function analyses.

ADHD, School Difficulties and No Diagnosis/Control. Discriminant function analyses were also performed using children in the ADHD plus School Difficulties group and the No Diagnosis Control group. Children in the ADHD plus School Difficulties group were subdivided into Predominately Inattentive Type, Predominately Hyperactive-Impulsive Type and Combined Type using the ADHD Rating Scale IV teacher form as previously described. Again, most of the children were assigned to the Combined group (N=30). The Predominantly Inattentive group consisted of 22 subjects and the No Diagnosis/Control group consisted of 23 subjects. The Hyperactive-Impulsive subtype was not included in the discriminant function analyses for the reasons listed previously. The total number of subjects included in the analysis was 75.

A discrete discriminant function was performed using omission scores on the CPT as the predictor of group status. Omission scores correctly classified 46.7% of the subjects according to initial group membership. This percentage is only slightly better than that expected by chance (33%). The percentage of subjects classified according to their initial group membership by CPT omission scores is presented in Table 11. Results suggest that the omission scores accurately classified 63.3% of the Combined Type group, 0% of the Predominately Inattentive Type group and 69.6% of the No Diagnosis/Control subjects. Interestingly, when omission scores were used as a predictor, none of the subjects were accurately classified or missclassified into the Predominately Inattentive Type group.

Another discriminant function analyses was conducted using the number of intervals during which subjects were engaged in off task behavior during the classroom observation as the predictor variable. The percentage of subjects correctly classified according to their initial group membership was 44% and was, again, not much better

Table 11

Cases Predicted by omissions scores on the Conners' Continuous Performance Test

		Predicted Group Membership		
Actual Group	# of cases	1	2	3
Group 1 (Combined type)	30	19 (63.3%)	0 (00.0%)	11 (36.7%)
Group 2 (Inattentive type)	22	12 (54.5%)	0 (00.0%)	10 (45.5%)
Group 3 (Control)	23	7 (30.4%)	0 (00.0%)	16 (69.6%)

Percent of cases correctly classified: 46.7%

Table 12

Cases Predicted by off task behavior during classroom observation

Actual Group	# of cases	Predicted Group Membership 1	2	3
Group 1 (Combined type)	30	26 (86.7%)	0 (00.0%)	4 (13.3%)
Group 2 (Inattentive type)	22	20 (90.9%)	0 (00.0%)	2 (9.1%)
Group 3 (Control)	23	16 (69.6%)	0 (00.0%)	7 (30.4%)

Percent of cases correctly classified: 44.0%

than chance (Table 12). Results suggest that the best prediction accuracy was found for the ADHD combined group.

Off task behavior accurately classified 86.7% of the ADHD Combined subjects into their original group with the remaining 13.3% being missclassified into the No Diagnosis/Control group. None of the Predominately Inattentive subjects were adequately classified into their original group. Almost all of these subjects (90.9%) were missclassified into the ADHD combined group. Only 30.4% of the control subjects were classified correctly according to their original group. Overall, predicted group membership resulted in no subjects being accurately classified or missclassified into the Predominantly Inattentive group.

When two CPT variables, omissions and variability SE scores, were used together as predictors, a discriminant function analyses correctly classified 52.0% of the subjects according their original group membership (Table 13). More specifically, the results suggest that the CPT scores together correctly classified 60% of the Combined subjects, 18.2% of the Predominately Inattentive subjects and 73.9% of the No Diagnosis/Control subjects.

Table 13

Cases Predicted by omission and commission scores on the Conners' Continuous Performance Test

		Predicted Group Membership		
Actual Group	# of cases	1	2	3
Group 1 (Combined type)	30	18 (60.0%)	3 (10.0%)	9 (30.0%)
Group 2 (Inattentive type)	22	11 (50.9%)	4 (18.2%)	7 (31.8%)
Group 3 (Control)	23	5 (21.7%)	1 (4.3%)	17 (73.9%)

Percent of cases correctly classified: 52.0%

Finally, the degree to which a combination of variables (e.g. CPT scores, classroom observation data, and CBM data) can predict membership into the identified groups was also investigated. Specifically, the variables, ADHD total behavior, total hits on the CPT, and words read correctly per minute during CBM probes were used as predictors when a discriminant function analyses was conducted. The resulting analyses indicated that these variables together correctly classified only 52.1% of the subjects into their original groups (Table 14).

Overall, results of the discriminant function analyses indicated that CPT variables, either alone or in combination, did not adequately predict membership defined groups (e.g. Combined, Predominantly Inattentive, or No Diagnosis/Control). Classroom observation data was also not able to predict original group membership to an acceptable degree. In summary, the discriminant function analyses indicated that none of the variables examined were good predictors of group membership.

Table 14

Cases Predicted by CPT scores, classroom behavior, and CBM data

Actual Group	# of cases	Predicted Group Membership 1	2	3
Group 1 (Combined type)	28	15 (53.6%)	7 (25.0%)	6 (21.4%)
Group 2 (Inattentive type)	22	9 (40.9%)	8 (36.4%)	5 (22.7%)
Group 3 (Control)	23	4 (17.4%)	4 (17.4%)	15 (65.2%)

Percent of cases correctly classified: 52.1%

DISCUSSION

Empirical support for the ecological validity of continuous performance tests has not yet been well established. Additionally, there is some literature which suggests these instruments are not useful for making diagnostic decisions. The primary goal of this study was to further investigate the ecological validity and the diagnostic utility of a continuous performance test. Specifically, the relationship between student scores on a Continuous Performance Test and student behavior in a classroom setting was examined. The ability of the Conners' Continuous Performance Test (CPT; Conners, 1992) to discriminate between children classified as normal and those classified as ADHD was also investigated.

Omission scores on the CPT, as well as, total hit scores and variability SE scores on the CPT were found to be significantly correlated with a number of classroom behaviors using a sample of children who exhibited ADHD characteristics in the school and home settings and normal controls. Correlations remained significant when a sample of children who exhibited ADHD characteristics primarily in the school setting was included. These findings are

consistent with those of other researchers and are similar to those reported by Barkley (1991) who used essentially the same observation coding system to observe children completing math problems while alone in a clinic room. Unlike the findings reported in the majority of other studies, CPT commission scores were not significantly correlated with any of the classroom behavior categories observed.

While the majority of correlations were significant, they were not very high. Disruptive behavior and CPT omission scores yielded the highest correlation (.548) indicating that the maximum amount of variance shared between the clinic based test scores and classroom observation data is 30%. Interestingly, the correlations between CPT and classroom behavior were slightly higher than the correlations between CPT scores and the number of math digits completed during the classroom observation; however, correlations between CPT data and behavior in the classroom were almost equivalent to correlations between CPT data and the number of words read correctly during a CBM probe. Because direct observations in the natural setting are considered by some to be an ideal criteria for

assessing the ecological validity of a laboratory measure and because correlating a laboratory measure with direct observations of the behavior of interest is one method used to establish the ecological validity of laboratory measures, it can be concluded that the correlations obtained during this study provide only limited evidence for the ecological validity of the CPT.

Results of a discriminate function analysis did not support the utility of the CPT for discriminating between children who exhibit characteristics of ADHD and normal controls. Neither CPT omission scores nor CPT commission scores adequately predicted group membership when teacher and parent rating scale data was used to identify groups. When the research sample was extended to include children who exhibited ADHD characteristics in the school setting only, CPT omission scores and CPT commission scores were again unable to adequately discriminate between groups. Overall, none of the CPT variables accurately predicted group membership alone or in combination with data from classroom observations and curriculum based measurement probes. These findings were consistent with other studies in which classification decisions based on CPT scores have

been discrepant with a diagnosis of ADHD based on parent interview and behavior rating scales (DuPaul, Anastopoulos, Shelton ,Guevremont, & Metevia, 1992).

There are several aspects of this study that are empirically unique. First, direct observation data was collected in the actual classroom setting of the research participants instead of in a simulated classroom or clinic playroom. This is important because classroom observations have been demonstrated to be better than analog observations for discriminating children with ADHD from normal controls (Platzman et al, 1992). It is also important because classroom observations provide information about behavior in the setting which is often most problematic for children with ADHD. Another unique aspect of this study is that all of the participants were observed while completing tasks designed to be within their instructional level. This allowed for better control of task difficulty across observations. The likelihood that off task behavior resulted primarily because of task difficulty was also decreased.

Overall, the results of this study suggest two important implications for clinical practice. First, CPT

scores and classroom behavior, although significantly related, are each unique. Thus, performance on the CPT should not be used to make inferences regarding behavior at school. Secondly, caution should be utilized when using the CPT to make diagnostic decisions. The predictive validity of this instrument appears to be low.

Several limitations to the present study should be recognized. First, this study utilized a relatively small sample size. A larger sample would have provided a more robust examination of the predictive accuracy of the CPT, classroom observation data and CBM probes. Additionally, a more in depth analysis of the predictive validity of the CPT using the various ADHD subtypes would also have been possible. Another limitation of this study is that rating scale data were used in isolation to assign participants to groups for the discriminant function analyses. Although, as Gordon (1993) has recognized, there is presently no independent "gold standard" for identifying ADHD, it is possible that the use of a more objective measure (e.g. classroom observation) would have provided a better assessment of ADHD characteristics and would have resulted in better differentiation between groups. A third

limitation of this study is that there was no assessment for the possibility of comorbid disorders among children in the ADHD group. The possible impact of comorbid disorders may have confounded the CPT performance and/or classroom behavior of this group. Finally, this study involved the participation of children who were referred through teacher nominations. It is possible that the utilization of a clinic-referred sample would have resulted in a group of children whose CPT responses and classroom behavior would have been more significantly different from that of typical peers.

Future research should focus on the replication and extension of current findings using other commercial CPT instruments, as well as, additional parameters of the CPT. Future research should involve the inclusion of clinic-referred children. The present study included teacher referred children who were not exhibiting any severe academic difficulties. It is likely that ADHD children who also possess academic difficulties represent a very different group. Finally, more research is necessary to investigate the actual construct that assessed using the CPT.

REFERENCES

Achenbach, T.M. & Edelbrock, C. (1983). Manual for the Child Behavior Checklist and Revised Child Behavior Profile. Burlington, VT: Thomas Achenbach.

Achenbach, T.M., & Edelbrock, C.S. (1986). Manual for the teacher report form and the child behavior profile. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T.M. & Edelbrock, C. (1991). Manual for the Child Behavior Checklist. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T.M., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. Psychological Bulletin, 101, 213-232.

Adams, J.A., & Humes, J.M. (1963). Monitoring of complex visual displays IV, training for vigilance. Human Factors, 5, 147-153.

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Anastasi, A. (1976). Psychological testing (4th ed.). New York: Macmillan.

Antonelli, D.C. and Karas, G.G. (1967). Performance and vigilance task under conditions of true and false knowledge of results. Perceptual and Motor Skills, 25, 129-138.

Baker, C.H. (1963). Further toward a theory of vigilance. In Vigilance: A symposium. New York: McGraw-Hill.

Barkley, R. (1991). Attention deficit hyperactivity disorder: A clinical workbook. New York: Guilford.

Barkley, R. (1990). Attention deficit hyperactivity disorder: A handbook of diagnosis and treatment. New York: Guilford Press.

Barkley, R.A., (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. Journal of Abnormal Child Psychology, 19, 149-178.

Barkley, R.A., & Cunningham, C.E. (1979). Stimulant drugs and activity level in hyperactive children. American Journal of Orthopsychiatry, 49, 491-499.

Barkley, R.A., DuPaul, G.J., & McMurray, M.B.(1990). A comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. Journal of Consulting and Clinical Psychology, 58, 775-789.

Barkley, R.A., DuPaul, G.J., & McMurray, M.B. (1991). Attention deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. Pediatrics, 87, 519-531.

Barkley, R.A., & Edelbrock, C.S. (1987). Assessing situational variation in children's problem behaviors: The Home and School Situations Questionnaires. In R.J. Prinz (Ed.), Advances in behavioral assessment of children and families (Vol. 3, pp.157-176). Greenwich, CT: JAI.

Barkley, R.A., Fischer, M., Newby, R.F, & Breen, M.J. (1988). Development of a multimethod clinical protocol for assessing stimulant drug response in children with attention deficit disorder. Journal of Clinical Child Psychology, 17(1), 14-24.

Breen, M.J. (1989). Cognitive and behavioral differences in AdHD boys and girls. Journal of Child Psychology and Psychiatry(5), 711-716.

Campbell, S. B., Douglas, V. L., and Morgenstern, G. (1971). Cognitive styles in hyperactive children and the

effect of methylphenidate. Journal of Child Psychology and Psychiatry,12, 55-67.

Campbell, C.K. & Rothschild, G.H. (1971). Cognitive styles in hyperactive children and the effect of methylphenidate. Journal of Child Psychology and Psychiatry, 12, 55-67.

Chee, P., Logan, G., Schachar, R., Lindsay, P., & Wachsmuth, R.(1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. Journal of abnormal Child Psychology, 17, 371-391.

Chee, P., Gordon, L., Schachar, R., Lindsay, P., & Wachsmuth, R. (1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. Journal of Abnormal Child Psychology, 17(4), 371-391.

Chinn, R.M. & Alluisi, E.A. (1964). Effect of three kinds of knowledge of results information on three measures of vigilance performance. Perceptual and Motor Skills, 18, 901-912.

Cofer, C.N.,and Appley, M.H. (1964). Motivation: theory and research. New York: Wiley.

Conners, C.K. (1969). A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 126, 884-888.

Conners, C.K. (1973). Rating scales for use in drug studies with children. Psychopharmacology Bulletin (Special Issue, Pharmacology of children), 24-29.

Conners, K.C. (1995). Conners' CPT computer program 3.0 user's manual. Multi-Health Systems.

Cooper, J.O. (1987). Planning and directing observational procedures. In J. Cooper, T. Heron, & W. Heward, Applied Behavior Analysis (pp.94-95). New York: MacMillan.

Corkum, P.V., & Siegel, L.S. (1993). Is the continuous performance task a valuable research tool for use with children with attention-deficit-hyperactivity disorder? Journal of Child Psychology and Psychiatry, 34(7), 1217-1239.

Davies, A.D., & Davies, D.R. (1975). The effects of noise and time of day upon age differences in performance at two checking tasks. Ergonomics, 18, 321-336.

Davies, D.R., & Parasuraman, R. (1982). The Psychology of Vigilance. London: Academic Press.

Davies, D.R. & Tune, G.S. (1969). Human vigilance Operformance. New York: American Elsevier.
Dember, W.N., & Warm, J.S. (1979). Psychology of Perception, 2nd edn. New York: Holt,

de Back, J. (1997). Behavior Evaluation Tool. Unpublished Manuscript.

Draeger, S., Prior, M., & Sanson, A. (1986). Visual and auditory attention performance in hyperactive children: Competence or compliance. Journal of Abnormal Child Psychology, 14(3), 411-424.

DuPaul, G.J. (1991). Parent and teacher ratings of ADHD symptoms: Psychometric properties in a community-based sample. Journal of Clinical Child Psychology, 20, 245-253.

DuPaul, G.J., Anastopoulos, A.D., Shelton, T.L., Guevremont, D.C., & Metevia, L. (1992). Multimethod assessment of attention- deficit hyperactivity disorder: The diagnostic utility of clinic-based tests. Journal of Clinical Child Psychology, 21(4), 394-402.

DuPaul, G.J., Anastopoulos, A.D., Power, T.J., Murphy, K., & Barkley, R.A. (1994). The ADHD Rating Scale-IV. Unpublished Manuscript, Lehigh University, Bethlehem, PA.

Egan, J.P., Greenberg, G.L., & Schulman, A.I. (1961). Operating characteristics, signal delectability, and the

method of free response. Journal of the Acoustical Society of America, 33, 993-1007.

Eliason, M.J., & Richman, L.C. (1987). The continuous performance test in learning and nonlearning disabled children. Journal of Learning Disabilities, 20, 614-619.

Fischer, M. Barkley, R.A., Edelbrock, C.S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria; II. Academic, attentional, and neuropsychological status. Journal for Consulting and Clinical Psychology, 5, 580-588.

Fischer, M., Newby, R., & Gordon, M. (1995). Who are the false negatives on continuous performance tests? Journal of Clinical Child Psychology, 24 (4), 427-433.

Gickling, E.E., & Rosenfield, S. (1995). Best Practices in Curriculum-Based Assessment. In A. Thomas & J. Grimes (Eds.), Best Practices in School Psychology (pp. 587-595). Washington, DC: The National Association of School Psychology.

Gordon, M. (1979). The assessment of impulsivity and mediating behaviors in hyperactive and nonhyperactive boys. Journal of Abnormal Child Psychology, 7(3), 317-326.
Gordon, M. (1983). The Gordon Diagnostic System. Boulder, CO:Clinical Diagnostic Systems.

Gordon, M. (1987). The Gordon Diagnostic System. DeWitt, NY:Gordon Systems, Inc.

Gordon, M., & Mettleman, B.B. (1988). The assessment of attention:I. Standardization and reliability of a behavior-based measure. Journal of Clinical Psychology, 44, 682-690.

Green, D.M., & Swets, J.A. (1966). Signal detection theory and psychophysics. New York: Wiley.

Gordon, M.,D., DiNiro, D., & Mettelman, B.B. (1988). Effect upon outcome of nuances in seletion criteria for ADHD/Hyperactivity. Psychological Reports, 62, 539-544.

Grant, M. I., Ilai,D., Nussbaum, N. L., & Bigler, E. D.(1990). The relationship between the continuous performance tasks and neuropsychological tests in children with attention-deficit hyperactivity disorder. Perceptual & Motor Skills, 70, 435-445.

Greenberg, L.M. (1987). An objective measure of methylphenidate response: Clinical use of the M.C.A. Psychopharmacology Bulletin, 23, 279-282.

Greenberg, L. (1989). Sensitivity of T.O.V.A. in undifferentiated attention deficit disorders, Unpublished manuscript.

Greenberg, L.M., & Waldman, I.D. (1993). Developmental normative data on the test of variables of attention (T.O.V.A.). Journal of Child Psychology and Psychiatry, 34(6), 1019-1030.

Griew, S., & Davies, D.R. (1962). The effect of aging on auditory vigilance performance. Journal of Gerontology, 17, 88-90.

Hair, J.F., Anderson, R.E., & Tatham, R.L. (1987). Multivariate Data Analysis: with readings. (2nd ed). New York: MacMillan.

Halcomb, C.G., & Kirk, R.E. (1965). Organismic variables as predictors of vigilance behavior. Perceptual and Motor Skills ,21, 547-552.

Halperin, J.M. (1991). The clinical assessment of attention. International Journal of Neuroscience, 58, 171-182.

Halperin, J.M., Sharma, V., Greenblatt, E., & Schwartz, S. (1991). Assessment of the continuous performance test: Reliability and validity in a nonreferred sample. Journal of Consulting and Clinical Psychology, 3, 603-608.

Halperin, J.F., Wolf, L.E., Pascualvaca, D.M., Newcorn, J.H., Healey, J.M., O'Brien, J.D., Morgenstern, A., & Young, J.G. (1988). Differential assessment of attention

and impulsivity in children. Journal of the American Academy of Child and Adolescent Psychiatry, 23, 285-290.

Hooks, K. Milich, R., & Lorch, E.P. (1994). Sustained and selective attention in boys with attention deficit hyperactivity disorder. Journal of Clinical Child Psychology, 23(1), 69-77.

Horn, W.F., Wagner, A.E., & Ialongo, N. (1989). Sex differences in school aged children with pervasive attention deficit hyperactivity disorder. Journal of Abnormal Child Psychology, 17, 109-125.

Kagan, J. (1964). The Matching Familiar Figures Test. Cambridge:Harvard University.

Kirby, N. H., Nettelback, T., & Bullock, J. (1978). Vigilance performance of mildly mentally retarded adults. American Journal of Mental Deficiency, 82, 394-397.

Kirby, N. H., Nettelback, T., & Bullock, J. (1979). Vigilance performance of mildly mentally retarded children. American Journal of Mental Deficiency, 84, 184-187.

Kirchner, G.L., & Knopf, I.J. (1974). Differences in the vigilance performance of second grade children as related to sex and achievement. Child Development, 45, 490-495.

Klee, S.H., & Garfinkel, B.D., (1983). The computerized continuous performance test: a new measure of inattention. Journal of Abnormal Child Psychology, 11, 487-496.

Knopf, I.J., & Mabel, R.M. (1975). Vigilance performance in second graders as a function of interstimulus intervals, socio-economic levels, and reading. Merrill-Palmer Quarterly,21, 195-203.

Koriath, U, Gualtieri, C.T, Van Bourgondien, M.E., Quade, D., & Werry, J.S. (1985). Construct validity of clinical diagnosis in pediatric psychiatry: Relationship

among measures. American Academy of Child Psychiatry, 24, 429-436.

Kupietz, S.S. (1976). Attentiveness in behaviorally deviant and nondeviant children-I. Auditory vigilance performance. Perceptual Motor Skills, 43, 1905-1101.

Kupietz, S.S., & Richardson, E. (1978). Children's vigilance performance and inattentiveness in the classroom. Journal of Child Psychology and Psychiatry, 19, 145-154.

Lassiter, K.S., D'Amato, C.D., Raggio, D.J., Whitten, J.C., & Bardos, A.N. (1994). The construct specificity of the continuous performance test: Does inattention relate to behavior and achievement? Developmental Neuropsychology, 10(2), 179-188.

Levy, F. (1980). The development of sustained attention (vigilance)and inhibition in children: some normative data. Journal of Child Psychology and Psychiatry, 21, 77-84.

Lindgren, S.K., & Lyons, D.A. (1984). Pediatric assessment of cognitive efficiency (PACE). Iowa City: University of Iowa, Department of Pediatrics.

Loeb, M., & Schmidt, E.A. (1960). Influence of time and task and false information on efficiency of responding to pure tones. U.S. Army Medical Research Laboratory Report. 426.

Lovejoy, M.C. & Rasmussen, N.H. (1990). The validity of vigilance tasks in differential diagnosis of children referred for attention and learning problems. Journal of Abnormal Child Psychology, 18, 671-681.

Luk, S.(1985). Direct observations studies of hyperactive behaviors. Journal of the American Child Psychology, 24, 338-344.

Mabel, R.M. (1968). The effect of instructional set and consequence of responding on vigilance in second and third

grade boys from two socio-economic levels. Doctoral Dissertation Emory University, 1968. Dissertation Abstracts International, 1969, 30, 850-851b.

Mackworth, J.F. (1950). Researches on the measurement of human performance. Medical Research Council, Special Report Series no. 268, 25-30, 74-78.

Mackworth, J.F. (1964). The effect of true and false knowledge of results on the detectability of signals in a vigilance task. Canadian Journal of Psychology, 18, 106-117.

Mackworth, J.F. (1970). Vigilance and attention. Maryland: Penguin.

Mackworth, N.H. (1957). Some factors affecting vigilance. Advancement of Science, 53, 389-393.

Margolis, J.S. (1972). Academic correlates of sustained attention. Unpublished doctoral dissertation. University of California at Los Angeles.

Margolis, J.S. (1973). Academic correlates of sus tained attention. In Warm, McClure, D.F., & Gordon, M. (1984). Performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12,(4), 561-572.

Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M.R. Shinn (Ed.), Curriculum-based measurement: Assessing special children (pp. 18-78). New York: Guilford.

Meents, C.K., (1989). Attention deficit disorder: A review of the literature. Psychology in the Schools, 26, 168-178.

Mesulam, G. (1985). In S. Goldstein & M. Goldstein, Managing attention disorders in children: A guide for practitioners. New York: Wiley.

McClure, F. D., & Gordon, M. (1984). The performance of disturbed hyperactive and nonhyperactive children on an objective measure of hyperactivity. Journal of Abnormal Child Psychology, 12, 561-572.

McCormack, P.D. (1959). Performance in a vigilance task with and without knowledge of results. Canadian Journal of Psychology(2), 68-72.

McGrath, J.J. (1960). The effects of irrelevant environmental stimulation on vigilance performance. In Warm, J.S.(Ed.), Sustained attention in human performance. New York: John Wiley & Sons Ltd.

Neal, G.L. & Pearson, R.G. (1966). Comparative effects of age, sex, and drugs upon tow tasks of auditory vigilance. Perceptual and Motor Skills, 23, 967-974.

Nuechterlein, K.H. (1983). Signal detection in vigilance tasks and behavioral attributes among offspring of schizophrenic mothers and among hyperactive children. Journal of Abnormal Psychology, 92, 4-28.

O'Dougherty, M., Nuechterlein, K., & Drew, B. (1984). Hyperactive and hypoxic Children: Signal Detection, Sustained Attention, and Behavior. Journal of Abnormal Psychology, 93, 178-191.

Parasuraman, R. (1976). Consistency of individual differences in human vigilance performance: an abilities classification analysis. Journal of Applied Psychology, 61, 486-492.

Parasuraman, R., & Davies, D.R. (1984). Varieties of attention. Toronto: Academic Press.

Pasculvaca, D.M., Wolf, L.E., Helaey, J.M., Tweedy, J.R., & Halperin, J.M. (1988, January). Sex differences in attention and behavior in school-age children. Paper presented at the meeting of the International Neuropsychological Society, New Orleans.

Pick, A.V., Frankel, D.G., & Hess, V.L. (1975). Children's attention: The development of selectivity. In Review of Child Development Research. Chicago: The University of Chicago Press.

Posner, M.I., & Boies, S.J (1971). Components of attention. Psychological Review, 78, 391-408.

Power, T.J. & DuPaul, G.J. (1996). Attention-Deficit Hyperactivity Disorder: the reemergence of subtypes. School Psychology Review, 25, 284-296.

Prinz, R.J., Tarnowski, K.J., & Nay, S.M. (1984). Assessment of sustained attention and distraction in children using a classroom analogue task. Journal of Clinical Child Psychology, 13(3), 250-256.

Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, ED., Jr & Beck, L. H. (1956). A continuous performance test of brain damage. Journal of Consulting Psychology, 20, 343-341.

Routh, D.K., Schroeder, C.S., & Tuama, L. (1974). Development of activity level in children. Developmental Psychology, 10, 163-168.

Rutschmann, J., Comblatt, B., & Erlenmeyer-Kimling, L. (1977). Sustained attention in children at risk for schizophrenia: Reports on a continuous performance test. Archives of General Psychiatry, 34, 571-575.

Sattler, J.M. (1990). Assessment of Children. San Diego: Jerome M. Sattler.

Schachar, R., Logan, G., Wachsmuth, R. & Chajezyk, D. (1988). Attaining and maintaining preparation: a comparison of attention in hyperactive, normal, and disturbed control group. Journal of Abnormal Child Psychology, 16, 361-378.

Schulman, J.L., & Reisman, J.M. (1959). An objective measure of hyperactivity. American Journal of Mental Deficiency, 64, 455-456.

Seidel, W.T. & Joschko, M. (1990) Evidence of difficulties in sustained attention in children with ADDH. Journal of Abnormal Child Psychology, 18(2), 217-229.

Sergeant, J. A. & Scholten, C.A. (1985). On resource strategy limitations in hyperactive: cognitive impulsivity reconsidered. Journal of Child Psychology and Psychiatry, 26, 97-109.

Shapiro, S.K., & Garfinkel, B.D. (1986). He occurrence of behavior disorders in children: The interdependence of attention deficit disorder and conduct disorder. Journal of the American Academy of Child Psychiatry, 25, 809-819.

Shinn, M.R. (1989) (Ed). Curriculum-based measurement: Assessing special children. New York: Guilford.

Sipowicz, R.R., Ware, R., & Baker, R. (1962). The effects of reward and knowledge of results on the performance of a simple vigilance task. Journal of Experimental Psychology, 64, 58-61.

Smith, I., Corkum, P., & Bryson, S. (1989). A total population study of aggression and hyperactivity II: Cognitive and social correlates. Unpublished manuscript, Dalhousie University, Halifax.

Solso, G. (1988). Cognitive Psychology. Old Tappan: Prentice Hall.

Sostek, A.J., Buchsbaum, M.S., & Rapoport, J.L. (1980). Effects of amphetamine on vigilance performance in normal and hyperactive children. Journal of Abnormal Child Psychology, 8, 491-500.

Swanson, H.L., & Cooney, J.B. (1989). Relationship between intelligence and vigilance in children. Journal of School Psychology, 27, 141-153.

Stanovich, K.E. Information processing in mentally retarded individuals. In Warm, J.S.(Ed.), Sustained attention in human performance. New York: John Wiley & Sons Ltd.

Sykes, D.H., Douglas, V.I., & Morgenstern, G. (1972). The effect of methylphenidate on sustained attention in hyperactive children. Psychopharmacologia (Berlin), 25, 262-274.

Sykes, D.H., Douglas, V.I., & Morgenstern, G. (1973). Sustained attention in hyperactive children. Child Psychology and Psychiatry, 14, 213-220.

Sykes, D.H., Douglas, V.I., Weiss, G., &, Minde, K.K. (1971). Attention in hyperactive children and the effect of Ritalin. Journal of Child Psychology and Psychiatry, 12, 129-139.

Talland, G.A. (1966). Visual signal detection as a function of age, input rate and signal frequency. Journal for Psychology,63, 105-115.

Tomporowski, P.D., & Allison, P. (1988). Sustained attention of adults with mental retardation. American Journal on Mental Retardation, 92(6), 531-538.

Trommer, B.L. Hoeppner, J.B., Lorber, R. & Armstrong, K. (1988). Pitfalls in the use of a continuous performance test as a diagnostic tool in attention deficit disorder. Developmental and Behavioral Pediatrics, 9(6), 339-345.

Tryon, W.W. (1984). Principles and methods of mechanically measuring motor activity. Behavioral Assessment, 6, 129-140.

Ullman, R.K., Sleator, E.K., & Sprague, R.L. (1985). A change of mind: The Conners abbreviated rating scales reconsidered. Journal of Abnormal Child Psychology, 13, 553-565.

van der Meere, J., & Sergeant, J. (1988). Controlled processing and vigilance in hyperactivity: time will tell. Journal of Abnormal Child Psychology, 16, 641-655.

Ware, J.R., Baker, R.A., & Sipowicz, R.R. Performance of mental deficient on a simple vigilance task. American Journal of Mental Deficiency, 66, 647-650.

Ware, J.R. (1961). Effects of intelligence on signal detection in visual and auditory monitoring. Perceptual and Motor Skills, 13, 99-102.

Warm, J.S. (Ed.)(1984). Sustained attention in Human Performance. New York: John Wiley & Sons Ltd.

Warm, J.S., Hagner, G.L., and Meyer, D. (1971). The partial reinforcement effect in a vigilance task. Perceptual and Motor Skills, 32, 987-993.

Warm, J.S., Kanfer, F.H., Kuwanda, S. and Clerk, J.L. Motivation in vigilance: effects of self-evaluation and experimenter-controlled feedback. Journal of Experimental Psychology, 92, 123-127.

Warm, J.S., Riechmann, S.W., Grasha, A.F., and Seibel, B. (1973). Motivation in vigilance: a test of the goal-setting hypothesis of the effectiveness of knowledge of results. Bulletin of the Psychonomic Society, 1, 291-292.

Weidenfeller, E.W., Baker, R.A., & Ware, J.R. (1962). Effects of knowledge of results (true and false) on vigilance performance. Perceptual Motor Skills, 14, 211-215.

Wiener, E.L. (1963). Knowledge of results and signal rate in monitoring: A transfer of training approach. Journal of Applied Psychology, 47, 214-222.

Wilkinson, R.T. (1961). Comparison of paced, unpaced, irregular and continuous displays in watchkeeping. Ergonomics, 4, 259- 267.

Werry, J.S., Elkind, G.S., & Reeves, J.C. (1987). Attention deficit, conduct, oppositional, and anxiety disorders in children: III. laboratory differences. Journal of Abnormal Child Psychology, 15, 409-428.

Wherry, J.N., Paal, N., Jolly, B.A., Holloway, B.E., & Vaught, L.(1993). Concurrent and discriminant validity of the Gordon diagnostic system: A preliminary study. Psychology in the Schools, 30, 29-36.

Wyatt, S., and Langdon, J.N.(1932). Inspection processes in industry. Industry Health Reasearch Board Roport No. 63. In Warm, J.S.(Ed.), Sustained attention in human performance. New York: John Wiley & Sons Ltd.

APPENDIX

BEHAVIORAL DEFINITIONS

OFF TASK - child breaks eye contact with his materials for more than 3 seconds

FIDGETING - any repetitive, purposeless motion of the legs, arms, hands, buttocks, or trunk. It must occur at least twice in succession to be considered repetitive, and it should serve no purpose. Examples include swaying back and forth, kicking one's legs back and forth, swinging arms at one's side, shuffling feet from side to side, shifting one's buttocks about in the chair, tapping a pencil or finger repeatedly on the table, and so on.

INAPPROPRIATE VOCALIZING - any vocal noise or verbalization made by the child that is clearly unrelated to the assignment. Examples: speech, whispering, singing, humming, making odd mouth noises, clicking one's teeth, and so on.

OUT OF SEAT - Any time the child's buttocks breaks contact with the flat surface of the seat.

PLAYING WITH OBJECTS - touching any object in the room besides the table, chair, or work materials. The child may touch his or her own clothing without being considered to play with an object. However, touching other objects in the room such as walls, light switches or blinds is coded in this category.

PEER ATTENTION - any one-to-one comment, smile, touch, or gesture directed to the target student by another child.

TEACHER ATTENTION - any one-on-one comment, smile, touch or gesture directed to the target student by the teacher.

VITA

Sandi Louise Spera born on June 6, 1968, in New Orleans, Louisiana. She received a bachelor of arts in psychology from the University of New Orleans in December of 1990. She received her masters degree in psychology at Louisiana State University in May of 1994. She is currently pursuing a doctor of philosophy in school psychology from Louisiana State University.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Sandi L. Spera

Major Field: Psychology

Title of Dissertation: Evaluation of a Continuous Performance Test

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination:

December 15, 1998

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989